Essays on the Functioning of Housing and Labor Markets

by

Christopher John Palmer

B.A. Economics; B.S. Mathematics, Brigham Young University (2008)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2014

Signature redacted

Author ....

Department of Economics

May 15, 2014

Signature redacted

Certified by ....

David Autor

Professor of Economics

Thesis Supervisor

Signature redacted

Certified by ....

Jerry Hausman

John and Jennie S. MacDonald Professor of Economics

Thesis Supervisor

Signature redacted

Certified by ....

Parag Pathak

Professor of Economics

Thesis Supervisor

Signature redacted

Certified by ....

William Wheaton

Professor of Economics

Thesis Supervisor

Signature redacted

Accepted by ....

Michael Greenstone

3M Professor of Economics

Chairman, Departmental Committee on Graduate Studies

Essays on the Functioning of Housing and Labor Markets

by

Christopher John Palmer

Submitted to the Department of Economics
on May 15, 2014, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The first chapter consists of my job-market paper. The foreclosure rate of subprime mortgages increased markedly across 2003–2007 borrower cohorts—subprime mortgages originated in 2006–2007 were roughly three times more likely to default within three years of origination than mortgages originated in 2003–2004. Many have argued that this surge in subprime defaults represents a deterioration in subprime lending standards over time. I quantify the importance of an alternative hypothesis: later cohorts defaulted at higher rates in large part because house price declines left them more likely to have negative equity. Using loan-level data, I find that changing borrower and loan characteristics explain approximately 30% of the difference in cohort default rates, with almost of all of the remaining heterogeneity across cohorts attributable to the price cycle. To account for the endogeneity of prices, I employ a nonlinear instrumental-variables approach that instruments for house price changes with long-run regional variation in house-price cyclicality. Control function results confirm that the relationship between price declines and defaults is causal and explains the majority of the disparity in cohort performance. I conclude that if 2006 borrowers had faced the same prices the average 2003 borrower did, their annual default rate would have dropped from 12% to 5.6%.

The second chapter is joint with David Autor and Parag Pathak. Externalities from the attributes and actions of neighborhood residents onto the value of surrounding properties and neighborhoods are central to the theory of urban economics and the development of efficient housing policy. This paper measures the capitalization of housing market externalities into residential housing values by studying the sudden and largely unanticipated 1995 elimination of stringent rent controls in Cambridge, Massachusetts, which had previously muted landlords' incentives to invest in their properties and altered the assignment of residents to locations. Pooling administrative data on the universe of assessed values and transacted prices of all Cambridge residential properties between 1988 and 2005, we find that rent decontrol generated substantial, robust price appreciation at decontrolled units and nearby never-controlled units, accounting for an estimated 30 percent of the $7.8 billion in Cambridge residential property appreciation during this period. The majority of this contribution is due to induced appreciation of never-controlled properties, while residential investments can explain only a small fraction of the total.

The third chapter is joint with Denis Chetverikov and Bradley Larsen. We present a methodology for estimating the distributional effects of an endogenous treatment that varies at the group level when there are group-level unobservables, a quantile extension of Hausman and Taylor (1981). Standard quantile regression techniques are inconsistent in this setting, even if the treatment is exogenous. Using the Bahadur representation of quantile estimators, we derive weak conditions on the growth of the number of observations per group that are sufficient for consistency and asymptotic normality. Simulations confirm the superiority of this grouped instrumental variables quantile regression estimator to standard quantile regression. An empirical application finds that low-wage earners in the U.S. from 1990–2007 were significantly more affected by increased Chinese import competition than high-wage earners. We also illustrate the usefulness of the estimation approach with additional empirical examples from urban economics, labor, regulation, and empirical auctions.

Chapter 1 Keywords: Mortgage Finance, Subprime Lending, Foreclosure Crisis, Negative Equity

Chapter 2 Keywords: Urban Economics, Residential Externalities, Rent Control, Price Regulations

Chapter 3 Keywords: Quantile Regression, Instrumental Variables, Panel Data, Wage Inequality, Import Competition

Chapter 1 JEL Classification: G01, G21, R31, R38

Chapter 2 JEL Classification: D61, H23, R23, R31, R32, R38

Chapter 3 JEL Classification: C21, C31, C33, C36, J30

Thesis Supervisor: David Autor
Title: Professor of Economics

Thesis Supervisor: Jerry Hausman
Title: John and Jennie S. MacDonald Professor of Economics

Thesis Supervisor: Parag Pathak
Title: Professor of Economics

Thesis Supervisor: William Wheaton
Title: Professor of Economics

Contents

Acknowledgments

Let me be the first to acknowledge that this (graduate school) wasn't easy. It was hard. Really hard. I felt deconstructed and reconstructed by the end. While I had many enjoyable moments, I also had some of the lowest moments of my life. My wife Megan stuck by me throughout, and I will never cease appreciating her for that. We had two children during graduate school, Quentin and Madeleine. They gave me perspective by being oblivious to the stresses and ups and downs of my program, exactly what my four amazing little sisters did for me growing up. Their joy and love for me means everything to me. My parents have been constants throughout graduate school. Their personal experience navigating an MIT PhD while raising a young family provided me with a tangible model that I remembered and tried to follow. They raised me to have energy, enthusiasm, ambition, correct priorities, and a genuine love for talking with people. Their prayers, love, and concern sustained me more than I will know.

My advisors, David Autor, Jerry Hausman, Parag Pathak, and Bill Wheaton, lived up to the MIT Economics billing. They were everything MIT was advertised to be: accessible, successful, engaged, wise. Coauthoring Chapter Two with David and Parag meant that I had a frequent reason to have close interaction with them, reaping advice, mentoring, and human capital. David has the uncanny ability to not lose sight of the forest while being willing to dig into the root system of the trees. As an advisor, his mastery of the unwritten social code of economics is invaluable. Jerry is a force of nature in his productivity, breadth of knowledge and experience, the quality of his Bayesian priors, and his insistence that I pursue econometric rigor in my work. Every one-on-one meeting I had with Parag felt like a concentrated dose of wisdom and insight, which benefited from a strong complementarity with his ability to impart advice without a hint of condescension, treating me as a peer. Bill Wheaton provided real estate expertise when I was most starved for it. He enabled me to have the subject credibility I needed when my interests led me in between the standard fields of economics.

My MIT classmates have been terrific. If I could buy stock in the future success (broadly defined) of my cohort, I would do so in a heartbeat. A steady stream of officemates made the day-to-day more palatable and inspired better habits in me. In particular, Danielle Li, Chris Walters, Isaiah Andrews, and Nils Wernerfelt took an interest in my work, functioned as public commitment devices when I struggled with procrastination or distraction, broke the monotony with their friendship and good humor, and tolerated frequent interruption to ask the kind of questions you wouldn't be comfortable asking anyone else.

Brad Larsen has been a brother-at-arms from day one. We overcame substantial odds in both getting admitted from BYU into MIT Economics in the same year. We shared a moving truck when we moved out here, and early morning, midday, and late-night conversations over problem sets, research, church duties, and kid play dates. Having such a loyal, believing, encouraging, and trusted friend buoyed me up and accelerated my own efforts.

John Arditi has occupied a special role for me at MIT. My life changed when his desk moved to be right outside mine in the last decaying moments of E52. He made sure I never got too drunk on the oft-egotistical Kool-Aid of academic economics. He has taken a genuine interest in my life and my children—they know his name. His friendship meant that I had someone to look forward to catching up with every day, and that was important to me.

Various institutions have been important facilitators of my graduate school experience. Without the National Science Foundation Graduate Research Fellowship, I would not have been admitted to MIT. Another NSF research grant for the research in Chapter Two provided our family with funding for several years that took the stressful edge off of financing graduate school. The Lincoln Institute and Institute for Humane Studies also helped in this regard. The Boston Federal Reserve provided no nominal funding, but data access without which I would not have been able to write Chapter One (or have the research agenda I now do).

I was blessed to attend an incredible alma mater institution. The BYU Economics department provided me with an abnormally large and involved set of mentors. I felt prepared for the culture and process of graduate school because of their close tutelage. When I was at my lowest in graduate school, their faith and perspective and own battle scars proved Balm of Gilead to me. Every return visit there is filled with the overwhelming feeling that I'm at my home turf with a friend in every office. The efforts of several merit explicit mention. The modal interaction I had with Jim McDonald felt like he was pouring love directly into my soul. Dave Sims took an early chance on me and set me on a research trajectory from which I have not departed that much. Being involved with his work on rent control opened up the opportunity for me to begin the research that led to Chapter Two of this thesis, with all of the fringe benefits (mentoring, funding, real estate bona fides, etc.) that came with that research. Frank McIntyre helped me charge forward on personal research on the returns to education in Armenia that proved pivotal in getting admitted to graduate school. James Cardon and Mark Showalter shared wit, advice, stories, and their own experience at several

key moments. Val Lambson made me feel like I belonged in academic economics, kept me from looking for the past of least resistance, and sold me on the value of economic theory.

A little-known group of BYU Economics alumni getting PhDs in Economics from Boston-area schools provided me with a monthly lunch fraternity. We mostly commiserated together, but also reminisced, advised, listened, lauded, encouraged, and ate. Those lunches will prove to be a very sweet memory of graduate school.

My first summer in college, I worked for Henry Pollakowski on a project studying Cambridge rent control, an experience that proved to have tremendous path dependence. That project set me on a trajectory that got me a job at an economics consulting company, which allowed me to earn enough to pay for college. Working for Henry gave me the background knowledge needed to work with Dave Sims, which, as I described above, has made a huge difference in my young career.

I have loved my time in my local Cambridge congregation of The Church of Jesus Christ of Latter-day Saints. Various responsibilities there have been time consuming but thrilling and rewarding. I have learned so much about faith, hope, doubt, and love from the people there. They challenged how I thought, and I am better for it. My most valued mentor there has been Chris Gillespie. The pep talks and accountability he provided me with down the stretch—emailing me in the weeks before my job-market paper was due to say, "Where is Section 3? Please email me Section 3 by 4 pm today or I will be on you."—were as important as any other inputs into writing my job-market paper.

Finally, I need to thank deity. It is a sensitive thing to thank God for success, lest those who feel unsuccessful infer that I credit God with their lack. And yet, I believe that God has sustained me throughout graduate school and that his hands are outstretched to those who seek them. The power of getting on my knees and humbly asking for help and strength to meet all of the demands of life head on has been ennobling and enabling. I vividly remember one prayer with Brad after a long study session in the dead of night before a 14.451 (macro theory) exam. I felt like Heavenly Father was listening and, while my struggles must seem dreadfully miniscule from a cosmic perspective, I felt that he understood the stress and anxiety I felt and was in my corner. I had hope that things would work out and that if they didn't, that would be OK, too. Crucially, the message of Christianity to me is that people can change, and this entire process has been one of change and improvement in me as a person. So in the end, thank God for the difficulty of graduate school—I know I am better for it in countless ways.

Chapter Acknowledgments

For Chapter One, I thank my advisors, David Autor, Jerry Hausman, Parag Pathak, and Bill Wheaton, for their feedback and encouragement; Isaiah Andrews, John Arditi, Matthew Baird, Neil Bhutta, Stan Carmack, Marco Di Maggio, Dan Fetter, Chris Foote, Chris Gillespie, Wills Hickman, Amir Kermani, Lauren Lambie-Hanson, Brad Larsen, Eric Lewis, Whitney Newey, Brian Palmer, Jonathan Parker, Bryan Perry, Jim Poterba, Brendan Price, Adam Sacarny, Dan Sullivan, Glenn Sueyoshi, Chris Walters, Nils Wernerfelt, Paul Willen, Heidi Williams, and Tyler Williams for helpful discussions and feedback; participants at the MIT Applied Microeconomics, Econometrics, Finance, and Public Finance workshops; and seminar participants at BYU, MIT, and Utah State. The loan-level data was provided by CoreLogic.

For Chapter Two, David Autor, Parag Pathak, and I thank seminar participants at Berkeley, the Boston Fed, Columbia, Harvard, MIT, NYU, Stanford, Virginia, and the NBER Summer Institute on Local Public Finance and Real Estate and our discussants, Erzo Luttmer and Jaren Pope, as well as the Editor of this journal and the referees for comments and suggestions that greatly improved the paper. We are grateful to David Sims for assistance with Cambridge Rent Control Board data and to Norma Coe, Cliff Cook, Bill Cunningham, Lisa Sweeney, and the staff at the Cambridge Assessor's Office for invaluable access to expertise and data. We acknowledge generous support from the Alfred P. Sloan Foundation, the Lincoln Institute for Land Policy, the National Science Foundation (grant SES-962572), and the Rappaport Institute for Greater Boston. I personally thank the National Science Foundation Graduate Research Fellowship (grant 0645960). We received excellent research assistance from Andrew Garin, Annalisa Scognamiglio, Karen Scott, Yuqi Song, Barrett Strickland, Daniel Sullivan, Thiago Vieira, Melanie Wasserman, and a hardworking team of MIT undergraduate data sleuths.

For Chapter Three, Denis Chetverikov, Brad Larsen and I thank Moshe Buchinsky, Ivan Canay, Brigham Frandsen, Antonio Galvao, Jin Hahn, Jerry Hausman, Rosa Matzkin, Whitney Newey, and Christopher Taber for helpful comments and Yuqi Song for meticulous research assistance.

Chapter 1

Why Did So Many Subprime Borrowers Default During the Crisis: Loose Credit or Plummeting Prices?

1.1 Introduction

Subprime residential mortgage loans were ground zero in the Great Recession, comprising over 50% of all 2006–2008 foreclosures despite the fact that only 13% of existing residential mortgages were subprime at the time.[1] The subprime default rate—the number of new subprime foreclosure starts as a fraction of outstanding subprime mortgages—tripled from under 6% in 2005 to 17% in 2009. By 2013, more than one in five subprime loans originated since 1995 had defaulted. While subprime borrowers by definition have been ex-ante judged as having greater default risk than non-subprime mortgages, many have pointed to the disproportionate growth in the share of defaults by subprime borrowers as evidence that the expansion in subprime lending was a major contributing cause to the housing crash of 2007–2009.

Why did the performance of subprime loans decline so sharply? A focal point of the discussion has been the stylized fact that subprime mortgages originated in 2005–2007 performed significantly worse than subprime mortgages originated in 2003–2004.[2] This is visible in the top panel of Figure

[1]Statistics derived from the Mortgage Bankers Association National Delinquency Survey. There is no standardized definition of a subprime mortgage, although the term always means a loan deemed to have elevated default risk. Popular classification methods include mortgages originated to borrowers with a credit score below certain thresholds, mortgages with an interest rate that exceeds the comparable Treasury Bill rate by 3%, certain mortgage product types, mortgages made by lenders who self-identify as making predominantly subprime mortgages, and mortgages serviced by firms that specialize in servicing subprime mortgages. For the purposes of this paper, subprime mortgages are defined as those in private-label mortgage-backed securities marketed as subprime, as in Mayer et al. (2009). For an estimate of the effects of foreclosures on the real economy, see Mian et al. (2011).

[2]See JEC (2007), Krugman (2007b), Gerardi et al. (2008), Haughwout et al. (2008), Mayer et al. (2009), Demyanyk

1, which uses data from subprime private-label mortgage-backed securities to show this pattern for 2003–2007 borrower cohorts.[3] Each line shows the fraction of borrowers in the indicated cohort that defaulted within a given number of months from origination.[4] The pronounced pattern is that the speed and frequency of default are higher for later cohorts—within any number of months since origination, more recent cohorts have defaulted at a higher rate (with the exception of the 2007 cohort in later years). For example, within two years of origination, approximately 20% of subprime mortgages originated in 2006–2007 had defaulted, in contrast with approximately 5% of 2003-vintage mortgages.

A popular explanation for the heterogeneity in cohort-level default rates over time is that loosening lending standards led to a change in the composition of subprime borrowers, potentially on both observable and unobservable dimensions (e.g. JEC, 2007 and COP, 2009). Others (e.g. Krugman, 2007a) blame an increase in the popularity of exotic mortgage products (for example, so-called balloon mortgages, which do not fully amortize over the mortgage term, leaving a substantial amount of principal due at maturity). The observed heterogeneity in cohort-level outcomes seen in Figure 1 could be generated by a decrease in the ex-ante creditworthiness of subprime borrowers over time or if the characteristics of originated mortgages became riskier. A third possibility is that price declines in the housing market—national prices declined by 37% between 2005–2009—differentially affected later cohorts, who had accumulated less equity when property values began to plummet. Being underwater—owing more on an asset than its current market value—could be an important friction in credit markets leading to a higher likelihood of default. Borrowers during a period of high price appreciation who have insufficient cash flow to make their mortgage payments can sell their homes or use their equity to refinance into a mortgage with a lower monthly payment. By contrast, if underwater homeowners cannot afford their mortgage payments, their alternatives are limited—lenders are often unwilling to refinance underwater mortgages or allow short sales (where the purchase price is insufficient to cover liens against the property).[5] The

and Van Hemert (2010), and Bhardwaj and Sengupta (2011) for examples of contrasting earlier and later borrower cohorts.

[3]This data will be discussed at length in Section 1.3. The analysis stops in 2007 because by 2008 the subprime securitized market was virtually nonexistent—the number of subprime loans originated in 2008 in the data fell by 99% from the number of 2007 originations.

[4]Following Sherlund (2008) and Mayer et al. (2009), I measure the point in time when a mortgage has defaulted as the first time that its delinquency status is marked as in foreclosure or real-estate owned provided it ultimately terminated without being paid off in full.

[5]Underwater homeowners may also default strategically to discharge their mortgage debt if they deem the option value of holding onto their property to be low. Bhutta et al. (2010) find that the property value of the median strategically defaulting borrower is less than half of the outstanding principal balance. Genesove and Mayer (1997) show that, all else equal, highly levered sellers also set higher reservation prices.

pattern of cohort default hazards could therefore come from four sources: price declines, changes in observable borrower characteristics, changes in unobservable borrower characteristics, and changes in mortgage product characteristics.

In this paper, I investigate the relative importance of each of these potential causes of declining cohort outcomes to understand what caused the increase in subprime defaults during the Great Recession. The counterfactual question I ask is whether 2003 borrowers (the best performing cohort) would have defaulted more like 2006 borrowers did if instead they had taken out mortgages in 2006 (when the worst performing cohort did). If so, then it is less plausible that deteriorating lending standards and risky mortgage products were a key driver of the surge in subprime defaults. On the other hand, if 2003 borrowers would have defaulted at a lower rate even after adjusting for observable borrower characteristics, loan characteristics, and market conditions, this would imply important differences in unobserved borrower quality across cohorts.[6]

To answer these questions, I estimate semiparametric hazard models of default using a panel of subprime loans that combines rich borrower and loan characteristics with monthly updates on loan balances, property values, delinquency statuses, and local price changes. I find that differential exposure to price declines explains 60% of the heterogeneity in cohort default rates. I also estimate that the product characteristics of subprime mortgages—but not the borrower characteristics—play an important role, accounting for 30% of the rise in defaults across cohorts. Conditioning on all three channels (price changes and loan and borrower characteristics) explains almost the entire change in cohort-level default rates, suggesting that the effect of any decline in unobserved borrower quality (e.g. from a deterioration in the accuracy of mortgage applications) was negligible. Returning to the counterfactual question posed above, my results imply that if 2006 borrowers had faced the prices that the average 2003 borrower did (i.e. at the same number of months since origination), 2006 borrowers would have had an annual default rate of 5.6% instead of 12%.[7] Furthermore, I find that if 2003 and 2006 borrowers had taken out identical mortgage products in addition to having faced the same prices, they would have defaulted at nearly identical rates.

House prices are an equilibrium outcome dependent on factors related to default risk. Whatever their source, price declines may have a causal effect on defaults. However, the potential for price

[6]Note that even absent a significant change in cohort quality, subprime lending could have had a sizable effect on the economy through feedback between subprime defaults and price declines. Isolating the causal effect of prices on defaults is thus an input into the larger question of what was the net impact of subprime lending on the housing market.

[7]I measure the annual default rate within five years of origination as 12 times the average fraction of existing loans that default each month.

changes and defaults to be caused by a third factor may lead to estimating a spurious relationship between price changes and defaults. In other words, some of the sources of price shocks may also have direct effects on the unobserved quality of borrowers and hence on defaults. A prominent hypothesis is that subprime penetration itself may subsequently have caused price declines and defaults, as suggested by Mayer and Sinai (2007), Mian and Sufi (2009), and Pavlov and Wachter (2009). In short, a credit expansion could amplify the price cycle, initially increasing prices from the positive demand shock as the pool of potential buyers grows. However, if the credit expansion involves a decrease in average borrower quality, this process will eventually lead to an increase in defaults, accelerating price declines. Thus, even though individual borrowers are price takers in the housing market, their unobserved quality may be correlated with the magnitude of price declines, resulting in biased estimates of the causal effect of prices on default risk.

The possibility of such a process makes it difficult to determine whether price changes actually cause defaults or if the defaults that are observed simultaneously with price declines are driven by the same latent factors driving prices and would therefore have occurred even absent any price changes. This impediment to estimating the causal effect of prices on defaults is also a challenge in estimating whether there were quality differences across cohorts. If unobserved quality differences affect both defaults and price declines, not taking into account the endogeneity of prices could lead to an underestimate of heterogeneity in cohort quality and an overestimate of the role of prices in affecting defaults.

To isolate the portion of cohort default rates driven by price changes from changes in unobserved borrower quality which also affect prices, I exploit plausibly exogenous long-run variation in metropolitan-area house-price cyclicality. As observed by Sinai (2012), there is persistence in the amplitude of house-price cycles—cities with strong price cycles in the 1980s were more likely to have strong cycles in the 2000s. I use this historical variation in house-price volatility to construct counterfactual price indices, which are unrelated to housing market shocks unique to the 2000s price cycle, e.g. because price volatility in the 1980s occurred well before the widespread adoption of subprime mortgages. Indeed, I show below that my instrument does not predict differential subprime expansion. Nevertheless, I also present evidence that areas that have cyclical housing markets also have cyclical labor markets. To address the possibility that price results could be explained by local labor shocks (an increase in the unemployment rate may cause defaults and depress prices), I verify that my results are robust to controlling for local unemployment rates.[8]

[8]Mayer (2010), Mian (2010), and Mian and Sufi (2012) argue that price declines first caused unemployment in

To my knowledge, this paper is the first to instrument for prices to address the joint endogeneity of prices and defaults in estimating the causal effect of price changes on defaults. While many researchers have looked at the relationship between house price appreciation and defaults, none of them have addressed the possible endogeneity of house price changes. For example, the common practice of imputing changes in property values using a metropolitan area home price index, although free from property-specific price shocks, does not address the concern that price changes at the metropolitan area level are themselves the outcome of demand and supply shocks that are likely correlated with unobserved borrower quality. Using a nonlinear instrumental variables approach to account for the endogeneity of covariates in a hazard model setting, I confirm that prices are endogenous, they are an important determinant of default, and they explain over half of the cohort pattern in default rates.

Figure 2 illustrates the differential effect that declining home prices had on origination cohorts by plotting the median mark-to-market combined loan-to-value ratio (CLTV) of each cohort of borrowers over time.[9] The beginning of each line shows the median CLTV at origination for mortgages taken out in January of that cohort's birth year. Thereafter, each line shows the median CLTV of all existing mortgages in the indicated origination cohort.[10] Each cohort's median CLTV began rising in 2007 as prices declined nationwide. However, there are two main differences between early and late cohorts. First, origination CLTVs increased over time—the median 2007 CLTV was 10 percentage points higher than the median 2003 CLTV, lending credence to the argument that underwriting standards deteriorated. Second, earlier cohorts' median CLTVs declined from origination until 2007 as prices rose (increasing the CLTV denominator) and as borrowers made their mortgage payments, reducing their indebtedness (the CLTV numerator), with the former effect dominating because of the low amount of principal paid off early in the mortgage amortization schedule. By contrast, later cohorts had not accumulated any appreciation or paid down any principal, as prices fell almost immediately after their origination dates. By early 2008, more than one-half of borrowers in both the 2006 and 2007 cohorts were underwater, and by early 2009, more than one-half of the 2005 cohort was underwater. Using variation in price changes across cities and cohorts and controlling for CLTV at origination, the empirical specifications below allow me to

the recent recession.

[9]The combined loan-to-value ratio (CLTV) of a mortgage is the sum of all outstanding principal balances secured by a given property divided by the value of that property. The data used in Figure 2 estimate market values from CoreLogic's Automated Valuation Model, see Section 1.3 for more details.

[10]Having a high CLTV at origination (equivalent to having a small down payment) is highly correlated with default risk and is routinely factored into the interest rates charged by lenders.

identify the causal effect of prices on defaults, differentiating between differences in negative equity prevalence across cohorts explained by high CLTVs at origination (a measure of cohort quality) and less opportunity to accumulate equity before price declines begin.

Suggestive evidence that the prevalence of negative equity affected economic outcomes is the bottom panel of Figure 1, which shows the cumulative prepayment probability by cohort—the fraction of each cohort's mortgages that had been paid off within the given number of months since origination.[11] The pattern across cohorts is exactly reversed from the cohort heterogeneity in default rates depicted in the top panel—more recent borrowers prepaid their mortgages much less frequently and at slower rates than borrowers from 2003–2005. Given the evidence that later cohorts were more likely to be underwater, the contrast between the cohort-level trends in defaults and prepayments is consistent with the notion that underwater borrowers in distress default and above-water borrowers in distress prepay.[12]

The differing experiences of the Pittsburgh and Minneapolis metropolitan areas serve as a motivating case study for the conceptual experiment in which this paper engages using geographic variation in prices. Although they had similar subprime market shares, these two cities had very different price cycles—Pittsburgh did not have much of a cycle, whereas Minneapolis home prices had a price cycle similar to the national average (see top panel of Figure 3).[13] As a consequence, the bottom panel of Figure 3 shows that the fraction of Pittsburgh subprime homeowners that were underwater stayed roughly constant at 30%, while the fraction of Minneapolis subprime homeowners who were underwater increased from under 20% before 2006 to over 35% by the middle of 2008.

The contrast between Pittsburgh and Minneapolis also extends to default rates. The top panel of Figure 4 shows that Pittsburgh cohorts defaulted at very similar rates, with later cohorts actually defaulting *less* than earlier cohorts by the end of the period. By comparison, in Minneapolis, where prices followed a boom-bust pattern, earlier borrower cohorts defaulted at a much lower rate than later cohorts. The bottom panel of Figure 4 shows that approximately 15% of Minneapolis subprime mortgages originated in 2006–2007 had defaulted within 12 months of origination, whereas only 5%

[11] Note that prepayment has a specific meaning in mortgage finance. As the issuer of a callable bond, a mortgage borrower has the prerogative to pay back the debt's principal balance at any time, releasing them of further obligation to the lender. In practice, this is done through refinancing or selling the home and using the proceeds to pay back the lender. See Mayer et al. (2010) for a discussion of mortgage prepayment penalties, an increasingly common feature of subprime mortgages.

[12] Note that this pattern could also be generated by cohort quality if riskier borrowers prepay less frequently, e.g. if they are less likely to trade-up to a more expensive home.

[13] According to Mayer and Pence (2008), 16% and 17% of mortgages originated in 2005 were subprime in Pittsburgh and Minneapolis, respectively.

of 2003–2004 mortgages had defaulted within the same time frame. The contrasting pattern across cohorts in Pittsburgh and Minneapolis suggests that the relative lack of a price decline and stable prevalence of negative equity in Pittsburgh may explain why default risk appears constant across Pittsburgh cohorts relative to Minneapolis, where an increasing share of underwater borrowers seems to have led to a rapid increase in default rates.

The strategy of this paper is to generalize the Pittsburgh-Minneapolis comparison to a comprehensive national dataset by including loan-level controls for the changing composition of borrowers in each locale and by isolating exogenous variation in each city's price cycle. Intuitively, I compare cohorts in areas with different price cycles (and thus different predicted availability of sell/refinance options for borrowers) to estimate whether they also had different default patterns after adjusting for observable underwriting characteristics.

There is a broad literature on the determinants of mortgage default.[14] A number of studies have examined the proximate causes of the subprime foreclosure crisis in particular (see Keys et al., 2008, Hubbard and Mayer, 2009, Mian and Sufi, 2009, and Dell'Ariccia et al., 2012). Kau et al. (2011) find that the market was aware of an ongoing decline in subprime borrower quality. Corbae and Quintin (2013) provide a model demonstrating how a period of relaxed underwriting standards could lead to a mass of mortgages originated to borrowers who would subsequently be extraordinarily sensitive to price declines.

Several papers have tried to quantify the relative contributions of underwriting standards and housing market conditions in the increase in the subprime default rate over time (all treating metropolitan area home price changes as exogenous) and have generally found a residual decrease in cohort quality. Sherlund (2008) concludes that leverage is the strongest predictor of increasing default risk and decreasing prepayment risk among subprime loans. Gerardi et al. (2008) use data through 2007 to ask whether lenders, investors, and rating agencies should have known that price declines would induce widespread defaults. Gerardi et al. (2007) examine the importance of negative equity. Krainer and Laderman (2011) examine the correlation between prepayment and default rates and find that declines in prepayment rates are strongly correlated with increases in default rates, particularly among borrowers with low credit scores. Bajari et al. (2008) estimate a dynamic model of default behavior on subprime mortgage data from 20 metropolitan areas and find evidence supporting both lending standards and price declines as drivers of default.

[14] For example, Deng et al. (2000), Foote et al. (2008), Pennington-Cross and Ho (2010), and Bhutta et al. (2010).

Other papers analyze differences in default or delinquency across cohorts. Mayer et al. (2009) demonstrate heterogeneity in the early default rates of origination cohorts and examine a series of bivariate correlations over time to document that loosening down payment requirements and declining home prices are both highly correlated with increases in early defaults. Bhardwaj and Sengupta (2012) estimate the cohort effects in default and prepayment hazards to be inversely related—later cohorts defaulted relatively more and prepaid relatively less. The most closely related study to this one is Demyanyk and Van Hemert (2011), which explicitly considers vintage effects in borrower quality and finds that prices played a much more important role than observable lending standards in explaining early delinquencies. Using data ending in 2008, they conclude that the bulk of the deterioration in cohort quality was due to unobservables, suggesting that the lending boom coincided with adverse selection among borrowers.

In summary, existing work has focused on whether changing underwriting standards (originated loan characteristics) explain changing default rates or whether prevailing market conditions such as negative equity were acute in areas where many borrowers are defaulting. They all find that a much larger portion of the deterioration in cohort quality is explained by home prices than ex-ante borrower characteristics. In contrast to these papers, with the benefit of several more years of data on the 2003–2007 subprime borrower cohorts and an instrumental-variables strategy, I am able to make causal inferences about the effect of price changes on default rates.

The paper proceeds as follows. Section 1.2 discusses the empirical strategy. I describe the data and compare the observable characteristics of borrower cohorts in Section 1.3. Identification concerns in the context of a hazard model are detailed in Section 1.4, along with a description of the estimator. After presenting initial descriptive estimates of the determinants of default that drive the cohort pattern, Section 1.5 presents the instrumental variables strategy and my main results, and Section 1.6 explores the economic mechanisms through which price declines affect default rates. Using my preferred empirical specification, I estimate cohort-level default rates under several counterfactual scenarios in Section 1.7. In Section 1.8, I conclude by summarizing my main findings and briefly discussing policy implications.

1.2 Empirical Strategy

Many factors determine default risk. Underwriting standards and market conditions, each predictive of future idiosyncratic income shocks and changes in prepayment opportunities, interact to generate defaults. Loose underwriting standards increase default rates because equally sized negative income shocks are more likely to prevent borrowers with high debt-to-income ratios from making mortgage payments and because borrowers with riskier income are more likely to have a negative shock that prevents them from making their mortgage payments. After a period of sustained price growth, younger loans are also relatively more sensitive to price declines because they have not accumulated as much equity and are thus more apt to be underwater and constrained in their ability to sell or refinance their mortgage. If an equal share of each cohort has an income shock that prohibits them from paying back their mortgage, cohorts with positive equity will simply sell their homes or refinance into mortgages with better terms. Later cohorts, on the other hand, have no such option and will default.

The objective of the hazard models presented below is to examine the relative importance of each of these factors by comparing loans with differing underwriting characteristics and in areas with differing price cycles to estimate how much of the heterogeneity in cohort default rates is explainable by each factor. Comparing observationally similar loans (i.e. by controlling for underwriting standards and loan age with a flexible baseline hazard specification) within a geography that were originated at different times allows me to take advantage of temporal variation in house prices within a geographic region. Likewise, comparing observationally similar loans taken out at the same time but in different cities utilizes spatial variation in house prices. To account for the endogeneity of the house price series of each geographic area, I estimate counterfactual price series by mapping each area's 1980–1995 house price volatility onto the most recent price cycle, as discussed in detail in Section 1.4 below. This setup allows me to decompose observed cohort heterogeneity into its driving factors by successively introducing additional controls that explain away the differences in cohort default rates.

1.2.1 Hazard Model Specification

I specify the origination-until-default duration as a proportional hazard model with time-varying covariates. Although the data are grouped into monthly observations, the proportional hazards functional form allows estimation of a continuous-time hazard model using discrete data (Prentice and Gloeckler, 1978 and Allison, 1982). Let the latent time-to-default random variable be denoted τ, and let the instantaneous probability (i.e. in continuous-time) of borrower i in cohort c and geography g defaulting at month t given that borrower i has not yet defaulted specified as

$$\lim_{\xi \to 0^+} \frac{\Pr\left(\tau \in (t-\xi,\, t] \middle| \tau > t-\xi\right)}{\xi} \equiv \lambda(X_{icg}(t),\, t) \tag{1.1}$$

$$= \exp(X'_{icg}(t)\beta)\lambda_0(t) \tag{1.2}$$

where $\lambda_0(\cdot)$ is the baseline hazard function that depends only on the time since origination t, and $X_{icg}(t)$ is a vector of time-varying covariates that in practice will be measured at discrete monthly intervals. The proportional hazards framework assumes that the conditional default probability depends on the elapsed duration only through a baseline hazard function that is shared by all mortgages. A convenience of this framework is that the coefficient vector β is readily interpretable as measuring the effect of the covariates on the log hazard rate.

Combining a nonparametric baseline hazard function with covariates entering through a parametric linear index function results in a semiparametric model of default. The specification for the covariates is

$$X'_{icg}(t)\beta = \gamma_c + W'_{B,i}\theta_B + W'_{L,i}\theta_L + \mu \cdot \Delta Prices_{icg}(t) + \alpha_g \tag{1.3}$$

where γ_c and α_g are cohort and geographic fixed effects, respectively; W_B and W_L are vectors of borrower (B) and loan (L) attributes, measured at the time of mortgage origination; and $\Delta Prices_{icg}(t)$ is a measure of the change in prices faced by property i at time t.[15] Borrower characteristics include the FICO score (a credit score measuring the quality of the borrower's credit history), debt-to-income (DTI) ratio (calculated using all outstanding debt obligations), an indicator variable for whether the borrower provided full documentation of income during underwriting, and an indicator variable for whether the property was to be occupied as a primary residence. Attributes of the mortgage note include the origination combined loan-to-value ratio (using all open liens on the property for the numerator and the sale price for the denominator), the mortgage interest rate, and indicator variables for adjustable-rate mortgages, cash-out refinance mortgages (when the new

[15]A natural concern with including fixed effects α_g in a nonlinear panel data model like this is the incidental parameters problem, which arises when the observations per group g is small and the number of groups grows with the sample size such that no progress is made in reducing the variance of the estimated fixed effects. Unlike a panel with fixed effects for each individual, the details of this application suggest this is not a significant worry. The number of observations per geography is already quite large, and as the total number of observations increases, the number of metropolitan areas in the U.S. remains fixed, leading to consistent estimates of α_g.

mortgage amount exceeds the outstanding principal due on the previous mortgage), mortgages with an interest-only period (when payments do not pay down any principal), balloon mortgages (non-fully amortizing mortgages that require a balloon payment at the end of the term), and mortgages accompanied by additional so-called piggyback mortgages.

The cohort fixed effects γ_c are the parameters of interest. As 2003 is the omitted cohort, the estimated baseline hazard function represents the conditional probability of default for a 2003 mortgage of each given age. The γ_c parameters scale this up or down depending on how cohort c mortgages default over their life-cycle, conditional on X and relative to 2003 mortgages of the same duration. Successively conditioning on geographic fixed effects, borrower characteristics, loan characteristics, and price changes reveals the extent to which each factor explains the systematic variation in default risk across cohorts. The estimated $\hat{\gamma}_c$ without conditioning on any covariates are a measure of the average performance of each cohort. Conditioning on prices, the γ_c are an estimate of the quality of each cohort, where quality is estimated using an ex-post measure (defaults). Conditioning on observable loan and borrower characteristics and prices, the γ_c represents the latent (i.e. unobserved) quality of each cohort. If cohort-level mortgage performance differences were driven by borrower unobservables, or if the explanatory power of the observables declined over time, then this would be captured by the cohort coefficients after controlling for all observables.

1.3 Data and Descriptive Statistics

In this section I briefly describe the data sources used in my analysis.

CoreLogic LoanPerformance (LP) Data. The main data source underlying this paper is the First American CoreLogic LoanPerformance (LP) Asset-Backed Securities database, a loan-level database providing detailed information on mortgages in private-label mortgage-backed securities including static borrower characteristics (DTI, FICO, owner-occupant, etc.), static loan characteristics (LTV, interest rate, purchase mortgage, etc.), and time-varying mortgage attributes updated monthly such as delinquency statuses and outstanding balances.[16] The LP data record monthly loan-level data on most private-label securitized mortgage balances, including an estimated 87%

[16] Using LP data is standard in the economics literature for microdata-based analysis of subprime and near-prime loan performance. See Sherlund (2008), Mayer et al. (2009), Demyanyk and Van Hemert (2011), and Krainer and Lederman (2011) for examples. See GAO (2010) for a more complete discussion of the LP database and comparison with other loan-level data sources.

coverage of outstanding subprime securitized balances. Because about 75% of 2001–2007 subprime mortgages were securitized, this results in over 65% coverage of the subprime mortgage market.[17] My estimation sample is formed from a 1% random sample of first-lien subprime mortgages originated in 2003–2007 in the LP database, resulting in a final dataset of over one million loan × month observations.[18]

Table 1 reports descriptive statistics for static (at time of origination) loan-level borrower and mortgage characteristics. On these observable dimensions, it is clear that subprime borrowers comprised a population with high ex-ante default risk.[19] The average subprime borrower in my data had a credit score of 617, slightly above the national 25th percentile FICO score and substantially below the national median score of 720 (Board of Governors of the Federal Reserve System, 2007). Among borrowers who reported their income on their mortgage application, the average back-end debt-to-income ratio, which combines monthly debt payments made to service all open property liens, was almost 40%, well above standard affordable housing thresholds. More than half of the loans in my estimation sample were for cash-out refinances, where the borrower is obtaining the new mortgage for an amount higher than the outstanding balance of the prior mortgage. As of April 2013, when my data end, 24% of the mortgages in my sample have defaulted and 50% have been paid off, leaving 26% of the loans in the data still outstanding.

Table 2 presents descriptive statistics by origination cohort. The distribution of many borrower characteristics is stable across cohorts. Average FICO scores, DTI ratios, combined loan-to-value ratios (measured using all concurrent mortgages and the sale price of the home, both at the time of origination), documentation status, and the fraction of loans that were owner-occupied or were taken out as part of a cash-out refinance are roughly constant across cohorts.[20] While there is substantial evidence that, pooling prime, near-prime, and subprime mortgages, borrower characteristics were deteriorating across cohorts (see JEC, 2007), the lack of a noticeable decrease in

[17] See Mayer and Pence (2008) for a description of the relative representativeness of subprime data sources. Foote et al. (2009) suggest that non-securitized subprime mortgages are less risky than securitized ones.

[18] As mentioned above, for my purposes a subprime loan is one that is in a mortgage-backed security that was marketed at issuance as subprime. I additionally drop mortgages originated for less than $10,000 and non-standard property types such as manufactured housing following Sherlund (2008).

[19] One measure of the elevated default risk inherent to subprime mortgages Gerardi et al. (2007), who find that homeownership experiences begun with a mortgage from a lender on the Department of Housing and Urban Development subprime lender list have a six times greater default hazard than ownership experiences that start with a prime mortgage.

[20] Note that the at-origination CLTVs reported here use the sale price of the home for its value, whereas the contemporaneous (mark-to-market) CLTVs in Figure 2 use estimated market values. If the divergence between these two measures over time is an important predictor of default, it will affect the magnitude of the estimated cohort main effects, which capture all unobserved factors changing across cohorts.

borrower observables in my data is consistent with observations from Gerardi et al. (2008) and Demyanyk and Van Hemert (2011) who argue that the declines within the population of subprime borrowers were too small to account for the heterogeneity in performance across cohorts.[21] Among mortgage product characteristics, however, there are important differences across cohorts, including a marked increase in prevalence of interest-only loans, mortgages with balloon payments, and mortgages accompanied by additional liens on the property. This finding of relatively stable borrower characteristics and large changes in certain mortgage characteristics is consistent with the findings of Mayer et al. (2009).

Specifications which directly examine the effects of negative equity make use of a novel feature of the LP dataset: contemporaneous combined loan-to-value ratios (CLTVs), which are a measure of the total amount of debt secured against a property relative to its market value. To calculate the CLTV numerator, CoreLogic uses public records filings on additional liens on the property to estimate the total debt secured against the property at origination. For the denominator, CoreLogic has an automated valuation model (popular in the mortgage lending industry) that uses the characteristics of a property combined with recent sales of comparable properties in the area and monthly home price indices to impute a value for each property in each month.

CoreLogic Home Price Index. For regional measures of home prices, I use the CoreLogic monthly Home Price Index (HPI) at the Core Based Statistical Area (CBSA) level.[22] These indices follow the Case-Shiller weighted repeat-sales methodology to construct a measure of quality-adjusted market prices from January 1976 to April 2013. They are available for several property categories—I use the single family combined index, which pools all single family structure types (condominiums, detached houses, etc.) and sale types (i.e. does not exclude distressed sales). Each CBSA's time series is normalized to 100 in January 2000.

The CoreLogic indices have distinct advantages over other widely used home price indices. The extensive geographic coverage (over 900 CBSAs) greatly exceeds the Case-Shiller index, which is only available for twenty metropolitan areas and the FHFA indices, which cover roughly 300 metropolitan areas. Unlike the FHFA home price series, CoreLogic HPIs are available for all residential

[21]Still, the nationwide decline in underwriting standards was driven in part by the subprime expansion: Even though the composition of the subprime borrower population was relatively stable over time, subprime borrowers represented a growing share of *overall* mortgage borrowers.

[22]There are 955 Core Based Statistical Areas in the United States, each of which is either a Metropolitan Statistical Area or a Micropolitan Statistical Area (a group of one or more counties with an urban core of 10,000–50,000 residents).

property types, not just conforming loans purchased by the GSEs. Finally, its historical coverage—dating back to 1976—predates the availability of deed-based data sources such as DataQuick that allow researchers to construct their own price indices but generally start only as early as 1988. I match loans to CBSAs using each loan's zip code, as provided by LP, and a 2008 crosswalk between zip codes and CBSAs available from the U.S. Census Bureau.[23]

Other Regional Data. For specifications that examine the importance of local labor market fluctuations, I use Metropolitan Statistical Area and Micropolitan Statistical Area unemployment rates from the Bureau of Labor Statistics (BLS) Local Area Unemployment Statistics series.[24] I also use publicly available Home Mortgage Disclosure Act (HMDA) data to calculate the subprime market share in a given CBSA $\times$ year by merging the lender IDs in the HMDA data with the Department of Housing and Urban Development subprime lender list as in Mian and Sufi (2009).[25] HMDA data discloses the census tract of each loan, which I allocate proportionally to CBSAs using a crosswalk from tracts to zip codes and then from zip codes to CBSAs.

1.4 Estimation and Identification

1.4.1 Estimation

Arranging the data into a monthly panel with a dependent variable $default_{icgt}$ equal to unity if existing mortgage i defaulted in month t, the likelihood $h(t)$ of observing failure for a given monthly observation must take into account the sample selection process. Namely, loans are not observed after they have defaulted, so the likelihood of sampling a given observation is a discrete hazard, which conditions on failure not having yet occurred. Suppressing dependence on X, the

[23] Available at http://www.census.gov/population/metro/data/other.html.

[24] Available at http://www.bls.gov/lau/home.htm.

[25] Using the HUD subprime lenders list to mark mortgages as subprime results in both false positives and false negatives: lenders who self-designate as predominantly subprime certainly issue prime mortgages as well, and non-subprime-identifying mortgage lenders also issue subprime mortgages. See Mayer and Pence (2008).

discrete hazard is

$$\begin{aligned}
h(t) &\equiv \Pr(default_{icgt} = 1) \\
&= \Pr(\tau \in (t-1,\, t] \big| \tau > t-1) \\
&= \int_{t-1}^{t} f(\tau)d\tau / S(t-1) \\
&= (F(t) - F(t-1))/S(t-1) \\
&= 1 - S(t)/S(t-1)
\end{aligned}$$

where $f(\cdot)$ and $F(\cdot)$ are the density and cumulative density of τ, the random variable representing mortgage duration until failure, and $S(\cdot) = 1 - F(\cdot)$ is the survivor function, the unconditional probability that observed mortgage duration exceeds the given amount of time. Using the familiar identity that $S(t) = \exp(-\Lambda(t))$, where $\Lambda(\cdot)$ is the integrated hazard function $\Lambda(t) = \int_0^t \lambda(\tau)d\tau$, I can rewrite the likelihood of observing failure for a given observation to be

$$\begin{aligned}
h(t|X) &= 1 - \exp(-\Lambda(t) + \Lambda(t-1)) \\
&= 1 - \exp\left(-\int_{t-1}^{t} \exp(X(\tau)'\beta)\lambda_0(\tau)d\tau\right)
\end{aligned}$$

where the last line used the specification of $\lambda(\cdot)$ in equation (1.2). If time-varying covariates are constant within each discrete time period (for example if the observed value of X_t represents the average of $X(\tau)$ for $\tau \in (t-1,\, t]$),

$$h(t|X) = 1 - \exp\left(-\exp(X_t'\beta)(\Lambda_0(t) - \Lambda_0(t-1))\right). \tag{1.4}$$

where $\Lambda_0(\cdot)$ is the integrated baseline hazard $\Lambda_0(t) = \int_0^t \lambda_0(\tau)d\tau$.

Incorporating this likelihood of observing $default_{icgt} = 1$, each month $\times$ loan observation's contribution to the overall log-likelihood is

$$\ell_{icgt} = default_{icgt} \cdot \log(h(t|X_{icgt})) + (1 - default_{icgt})\log(1 - h(t|X_{icgt})). \tag{1.5}$$

I can then estimate the hazard model parameters of equation (1.2) by Quasi-Maximum Likelihood (MLE) in a Generalized Linear Model framework where the link function $G(\cdot)$ satisfying $h(t) =$

$G^{-1}(X_t'\beta + \psi_t)$ is the complementary log-log function

$$G(h(t)) = \log(-\log(1 - h(t))) = X_t'\beta + \underbrace{\log(\Lambda_0(t) - \Lambda_0(t-1))}_{\psi_t}.$$

Estimating a full set of dummies ψ_t allows for the baseline hazard to be fully nonparametric à la Han and Hausman (1990).[26] The estimates of the baseline hazard function represent the average value of the continuous-time baseline hazard function $\lambda_0(\cdot)$ over each discrete interval $\bar{\lambda}_{0t} = \int_{t-1}^{t} \lambda_0(\tau)d\tau$ and are obtained as $\hat{\bar{\lambda}}_{0t} = \exp(\hat{\psi}_t)$. Under the usual MLE regularity conditions, estimates of β and ψ will be consistent and asymptotically normal.

1.4.2 Identification

The proportional hazard model is identified—implying that the population objective function is uniquely maximized at the true parameter values—under the assumptions that 1) conditional on current covariates, past and future covariates do not enter the hazard (often termed strict exogeneity), and 2) any sample attrition is unrelated to the covariates (Wooldridge, 2007).[27] Stated in terms of the conditional distribution $F(\cdot|\cdot)$ of failure times τ, the strict exogeneity and non-informative censoring assumptions are met provided

$$F\left(\tau | \tau > t-1, \{X_{icgs}, c_{is}\}_{s=1}^{T}\right) = F(\tau|\tau > t-1, X_{icgt})$$

where c_{is} is an indicator for whether loan i was censored at time s. In principle, if lags or leads of the covariates enter into λ, the strict exogeneity condition can be satisfied by including them as explanatory variables in the vector X_{icgt}.

An important form of censoring in mortgage data arises from borrowers paying back their mortgages in full. Mortgages that have been prepaid are treated as censored because all that can be learned about their latent time until termination by default is that it is at least as long as the observed elapsed time until prepayment. Technically, any such hazard model with multiple failure types is a competing risks model, which can be generalized to accommodate the potential dependence of one

[26] Alternatively, ψ_t can be thought of as estimating a piecewise-constant baseline hazard function. As discussed above in the context of the geographic fixed effects, the incidental parameters problem is not a concern here since increases in sample size (the number of loans) would not increase the number of ψ needing to be estimated.

[27] The linear-index functional form assumption that the effect of covariates on the hazard is linear in logs is not necessary for identification and is made for the sake of parsimony and convenience in interpreting the coefficients.

risk on shocks to another. Under the assumption there is no unobserved individual heterogeneity in the default hazard (or that unobserved heterogeneity in the default and prepayment hazards are independent at the individual level), competing risks models can be estimated as separable hazard models with observations representing other failure types treated as censored.[28] As in Gerardi et al. (2008), Sherlund (2008), Foote et al. (2010), and Demyanyk and Van Hemert (2011), I adopt this approach and focus on estimation of the default hazard.[29] I also verify the robustness of my main results to allowing for unobserved heterogeneity in the default hazard.

Turning to causality, the key identifying assumption for the estimated coefficient μ in equation (1.3) to be interpretable as the causal effect of the decline in property values is that fluctuations in home prices and unobserved shocks to default risk are independent.[30] To illustrate how the exogeneity of X affects estimates of β in a hazard model setting, consider the case of time-invariant covariates and no censoring. In this simplified setting, the exogeneity condition necessary for the maximum likelihood estimates of the hazard model parameters to represent causal effects is that the probability of failure (conditional on reaching a given period) is correctly specified in (1.2) and (1.3). Again, letting τ be the random variable denoting the mortgage duration until failure, the formal condition is

$$\lim_{\xi \to 0^+} E\left[\frac{1\left(\tau \in (t-\xi, t]\right)}{\xi} - \lambda(X_{icgt}, t)\middle| X, \tau > t - \xi\right] = 0 \tag{1.6}$$

where $1(\cdot)$ is the indicator function. Analogous to omitted variables bias in a linear regression, this condition would be violated if there were an omitted factor ω which affects default rates and is not independent of X. In this case, misspecification leads to violation of the exogeneity assumption because ω affects failure, is not in λ, and survives conditioning on X. To see this, suppose that the true instantaneous probability of default conditional on $\tau > t - \xi$ is not $\lambda(X, t)$ but is

$$\tilde{\lambda}(X, \omega, t) = \exp(X\beta + \omega)\tilde{\lambda}_0(t),$$

[28] See Heckman and Honoré (1989) for a full discussion of identification in competing risks models.

[29] The most well-known example of allowing for correlated default and prepayment unobserved heterogeneity is Deng et al. (2000), who jointly estimate a competing risks model of mortgage termination using the mass-points estimator of McCall (1996).

[30] This condition is stronger than price and default shocks being uncorrelated and is required in non-additive models. See Imbens (2007) for a discussion.

where both X and ω may depend on t. Then

$$\begin{aligned}\lim_{\xi\to 0^+} E\left[\frac{1\left(t-\xi<\tau\leq t\right])}{\xi} - \lambda(X,t)\Big|X,\tau>t-\xi\right] &= E\left[\tilde{\lambda}(X,\omega,t)\big|X\right] - \lambda(X,t) \\ &= E\left[e^{\omega}\exp(X\beta)\tilde{\lambda}_0(t)\big|X\right] - \lambda(X,t).\end{aligned}$$

If ω and X are independent, then the exogeneity condition becomes

$$E\left[e^{\omega}\exp(X\beta)\tilde{\lambda}_0(t)\big|X\right] - \lambda(X,t) = \exp(X\beta)E\left[e^{\omega}\right]\tilde{\lambda}_0(t) - \exp(X\beta)\lambda_0(t).$$

Thus, the presence of independent ω simply scales the estimate of the baseline hazard function. In other words, the baseline hazard function estimated without controlling for ω will be estimating $E\left[e^{\omega}\right]\tilde{\lambda}_0(t)$—but the estimation of the slope coefficients will be unaffected and the exogeneity condition of equation (1.6) will hold in expectation. However, if ω and X are not independent, then the omission of ω leads to a violation of equation (1.6), and estimated β will not represent the marginal effect of X on default, as discussed in Section 1.5.2 below.[31]

In the general case, even independent unobserved heterogeneity will affect the conditional distribution of $\tau|X$ (and hence the estimated coefficients), a common obstacle in nonlinear panel models. Lancaster (1979) introduced the Mixed Proportional Hazard (MPH) model where the heterogeneity enters in multiplicatively (additively in logs).[32] Conditional on unobserved heterogeneity ε, the hazard function becomes

$$\lambda(t|X_{icgt},\varepsilon_i) = \exp(X'_{icgt}\beta + \varepsilon_i)\lambda_0(t). \tag{1.7}$$

The literature on unobserved heterogeneity in duration models has broadly found that ignoring unobserved heterogeneity biases estimated coefficients down in magnitude. Intuitively, the presence of ε induces survivorship bias—loans with low draws of ε last longer and are thus overrepresented in the sample relative to their observables. Individuals whose observable characteristics put them at a high ex-ante risk of default and yet have lengthy durations are likely observed in the sample because they have low unobserved individual-specific default risk (high latent quality). The negative correlation between X and ε induced by the sample selection process can prevent consistent

[31]Estimating a proportional hazard model with no censoring and time-invariant covariates is equivalent to a linear regression of log duration on the covariates (Wooldridge, 2007). This illustrates why this special case permits unobserved heterogeneity provided it is independent of the covariates; in a linear model, additive unobservables affect the consistency of the parameter estimates only if they are correlated with the covariates.

[32]Elbers and Ridder (1982) showed that the MPH model is identified provided there is at least minimal variation in the regressors.

estimation of β.

Equation (1.7) pins down the conditional distribution F of latent failure times τ to be

$$F(\tau|X_{icgt}, \varepsilon_i) = 1 - \exp\left(-\Lambda((t|X_{icgt}, \varepsilon_i)\right)$$

where $\Lambda(\cdot|X, \varepsilon)$ is the integrated hazard. Specifying the distribution of ε to have cumulative distribution function $G(\cdot)$, the distribution $\tilde{F}(\tau|X_{icgt})$ of $\tau|X$ is then obtained by integrating out ε:

$$\tilde{F}(\tau|X_{icgt}) = \int_{-\infty}^{\infty} F(\tau|X_{icgt}, \varepsilon_i) dG(\varepsilon_i).$$

Finally, the modified likelihood $\tilde{h}(t|X)$ of observing failure at time $\tau \in (t-1,\, t]$ is

$$\tilde{h}(t|X) = 1 - \tilde{S}(t|X)/\tilde{S}(t-1|X) \tag{1.8}$$

where the new survivor function is denoted $\tilde{S}(\cdot|X) = 1 - \tilde{F}(\cdot|X)$. Estimation then proceeds by replacing $h(\cdot|X)$ with $\tilde{h}(\cdot|X)$ in the log-likelihood expression of equation (1.5). After presenting my main results, I verify that my results are robust to the presence of independent unobserved heterogeneity by specifying $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ so that $G(\varepsilon) = \Phi(\varepsilon/\sigma)$, where $\Phi(\cdot)$ is the standard normal cumulative density function.[33]

1.4.3 Isolating Long-Run Variation in Housing Price Cycles

One example of an omitted factor that may be correlated with X is the expansion of subprime credit, which may initially increase prices as a positive shock to the demand for owner-occupied housing, as suggested by Mayer and Sinai (2007), Mian and Sufi (2009), and Pavlov and Wachter (2009). If the credit expansion leads to a decrease in the quality of the marginal borrower, prices will eventually fall as these riskier borrowers default, depressing prices both from a positive shock to the supply of owner-occupied housing on the market and from negative foreclosure externalities (see Hartley, 2010 and Campbell et al., 2011).[34] Thus, the expansion of subprime credit may be

[33]There is a large literature on the relative merits of parametric assumptions on the baseline hazard function and the unobserved heterogeneity distribution. See Lancaster (1979), Heckman and Singer (1984), Han and Hausman (1990), Meyer (1992), Horowitz (1999), and Hausman and Woutersen (2012).

[34]Dagher and Fu (2011) provide an example of the mechanism behind such an expansion: counties that had significant entry of non-bank mortgage lenders had stronger growth in credit and prices, as well as stronger subsequent increases in defaults and decreases in prices. Brueckner et al. (2012) offer a model of how price increases could fuel lender expectations and further credit expansion. Berger and Udell (2004) also discuss empirical evidence of underwriting standards deteriorating during a credit expansion.

an omitted variable that directly affects both defaults (by decreasing the quality of the marginal subprime borrower) and prices, potentially leading to a spurious estimated relationship between prices and defaults. A related worry from the perspective of the exogeneity condition in equation (1.6) is that areas with the strongest price declines are also likely the areas hit hardest by the recession. If a negative employment shock simultaneously causes both defaults and price declines, then local labor market strength may be an important omitted variable that biases the estimates towards finding an effect of prices on default. Below, I discuss how I account for each of these potential biases.

To address these endogeneity concerns, I develop an instrument that isolates the long-run component of each Core Based Statistical Area's (CBSA) price cycle and is arguably independent of contemporaneous shocks to prices or default rates, e.g. from credit or labor market fluctuations. The CoreLogic repeat-sales price index for each CBSA, discussed in greater detail above, provide a measure of the relative level of nominal house prices in a given CBSA $\times$ month, denoted here as HPI_{gt}. Sinai (2012) notes that a similar set of metropolitan areas had large 1980s and 2000s price cycles. Using this persistence, I determine the portion of a CBSA's price cycle that is predictable using only the historical cyclicality of that city. First, I form a summary measure σ_g^P quantifying the long-run cyclicality of CBSA g defined as the standard deviation of monthly changes in the CoreLogic repeat sales home price index from 1980-1995

$$\sigma_g^P \equiv \left(\frac{1}{T-1}\sum_{t\in\mathcal{T}}(\Delta HPI_{gt} - \overline{\Delta HPI}_g)^2\right)^{1/2} \tag{1.9}$$

where $T = 180$ is the number of months over which the standard deviation is calculated; $\mathcal{T}$ is the set of months from January 1980 to December 1995, inclusive; $\Delta HPI_{gt} = HPI_{gt} - HPI_{gt-1}$; and $\overline{\Delta HPI}_g$ is the average value of ΔHPI_{gt} for CBSA g and $t \in \mathcal{T}$.[35] Figure 5 shows the average value of the CoreLogic repeat sales home price index by quartile of σ^P. The persistence in price volatility isolated by the first stage is visible: the average price cycle in the late 2000s was much more pronounced for CBSAs that had stronger price cycles in the 1980s, that is, higher quartiles of σ^P have monotonically stronger price cycles.

[35] I calculate the standard deviation of the first differences in the HPI variable to emphasize the importance of the (low-frequency) price cycle. CBSAs with high variance of HPI in levels (as opposed to high ΔHPI) could simply be areas that had sustained price growth or high-frequency volatility.

1.5 Results

1.5.1 Results Treating Price Changes as Exogenous

Table 3 reports estimates of equation (1.2) using the estimator described above, treating price changes as exogenous to offer initial estimates of the relationship between price changes, underwriting standards, and cohort-level differences in default rates. I cluster all standard errors at the CBSA level to account for area-specific shocks to the default rate in inference. All specifications include nonparametric controls for the baseline hazard function.[36] Column 1 includes only cohort fixed effects to quantify the pattern of declining cohort-level performance from Figure 1 in a hazard-model framework. These coefficients can be interpreted as the change in the log hazard rate and imply, for example, that subprime loans in the 2007 cohort had a default hazard 73 log points greater than the 2003 cohort (the omitted category). These unadjusted cohort coefficients are large and precisely estimated, implying that the probability of a 2005–2007 cohort mortgage defaulting in any given month conditional on the mortgage having survived to that month is more than twice as high as 2003 cohort mortgages. Column 2 adds fixed effects for each CBSA in the sample (570 fixed effects) to verify that cohort differences are not driven by the geographic composition of later cohorts. Conditioning on CBSA fixed effects does not materially affect the estimated differences in cohort default hazards.

Columns 3 and 4 add borrower characteristics and loan characteristics, respectively, as detailed in Section 1.3. The coefficients on these credit risk factors all have intuitive signs. Borrowers had higher default rates if they lacked full income documentation, were not owner-occupants, or had lower FICO scores and higher DTI ratios. Mortgages defaulted more frequently if they were non-fixed rate mortgages, had higher CLTVs or interest rates, or were accompanied by additional liens. Column 3, which includes only borrower characteristics, shows that the adjusted default hazard of earlier cohorts is higher than in column 1, suggesting that, relative to 2003 borrowers, 2004 and 2005 subprime borrowers underperformed relative to what would be expected based on their individual attributes. For 2006–2007 cohorts, the differential default hazard is lower than in column 1, although the average decrease between column 1 cohort effects and column 3 cohort effects is approximately zero. The inability of borrower characteristics to substantively explain the

[36]The baseline hazard controls consist of an indicator variable for each possible value of loan age from 1–70 months, with the final indicator variable also turned on for all values of loan age exceeding 70 months. The estimated baseline hazard functions resemble the hump-shaped baseline hazards of Deng et al. (2000) and are available from the author on request.

cohort-level differences is not surprising given the summary statistics reviewed above showing that the mean observable attributes of borrowers are not changing much across cohorts.[37] The results of column 4 tell a different story: including controls for loan characteristics and not borrower characteristics explains on average 24% of the unadjusted cohort effects estimated in column 1. This suggests that the loan characteristics that were changing across cohorts were an important driver of defaults. Conditioning on both borrower and loan characteristics together in column 5 reduces the residual cohort heterogeneity (i.e. the column 4 coefficients relative to the column 1 coefficients) by an average of 29%.

To get a sense of which covariates are most important in explaining the cohort pattern, I estimated the specification of column 5, leaving out one characteristic at a time. Three characteristics stand out as contributing substantially the attenuation of the estimated cohort effects: the balloon and interest-only dummies and the loan interest rate. As the interest rate should represent everything that the market knew about the riskiness of the loan, its importance reenforces that priced observables are important in predicting the cohort-level default pattern. The importance of the balloon and interest-only indicators is consistent with Table 2, which showed that balloon mortgages and interest-only mortgages were the two product characteristics that changed the most across cohorts and thus had the strongest potential to explain cohort-level defaults.

Column 6 drops all borrower- and loan-level covariates and instead controls for the 12-month change in log of the CoreLogic repeat-sales Home Price Index (HPI), defined at the CBSA-level as

$$\Delta \log(HPI_{icgt}) \equiv \log(HPI_{icgt}) - \log(HPI_{icgt-12}) \tag{1.10}$$

where HPI_{icgt} is the value of the CoreLogic repeat-sales price index for CBSA g in the calendar month corresponding to loan i having a duration of t.[38] This variable is a strong predictor of default. The coefficient on the 12-month change in log HPI implies that properties experiencing the 75th percentile 12-month price change (+5%) would have a 33% lower hazard than properties exposed to the 25th percentile 12-month change in prices (–5%), corresponding to an approximately one percentage point decrease in the annual default rate. Controlling for the 12-month change in prices,

[37]While individual borrower characteristics do not explain much of the differences in default rates across cohorts, they are individually strong predictors of default, as evidenced by the large increase in the log likelihood value between columns 2 and 3.

[38]I index HPI by i as well to emphasize that in my notation t refers to event time (i.e. loan age). Even though HPI only varies by CBSA × calendar month, for example, not all six-month old ($t = 6$) mortgages in CBSA g have the same HPI value.

the cohort effects in column 6 are lower than the estimates in column 5, showing that price changes in the most recent 12 months seem to be more closely related to observed cohort heterogeneity than borrower and loan characteristics. The residual differences in default rates across cohorts decrease on average by 50% (depending on the cohort) relative to the baseline cohort coefficients in column 1.

Controlling for both borrower and loan characteristics and price changes leaves little cohort-level heterogeneity unexplained. The estimates in column 7 of the latent quality of each cohort (i.e. the portion of cohort outcomes not attributable to price changes or individual-level controls) are statistically insignificant with the exception of the 2005 cohort. While statistically significant, more than 70% of the unadjusted estimate of the difference between the 2003 and 2005 cohorts (column 1) is explained by prices and observables.

These results illustrate that observable loan characteristics and prices play important roles in explaining the heterogeneity in default rates across origination cohorts, together explaining on average 95% of the cohort disparities in column 1.[39] In particular, places where price declines are greater have higher default rates, and the incidence of these price declines is disproportionately borne by later cohorts. I now turn towards developing causal estimates of the impact of prices on default behavior.

1.5.1.1 Unobserved Heterogeneity

This section examines the robustness of the above results to misspecification from ignoring independent unobserved heterogeneity ε by allowing the true hazard model to be specified as in (1.7). The results of maximizing the sample log-likelihood function described by (1.5), replacing $h(t|X)$ with $\tilde{h}(t|X)$ defined in equation (1.8) and modeling $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, are presented in Table 4. There are two important caveats in comparing these results to the results of Table 3. Because of the computational burden of maximizing the likelihood while integrating out the unobserved Gaussian heterogeneity, columns 1–4 do not include geographic fixed effects or cluster standard errors by CBSA as in the rest of the paper. Column 5 includes state fixed effects to test how sensitive the

[39]The additional explanatory power gained from controlling for prices and characteristics simultaneously suggests that there are important interactions between prices and loan and borrower characteristics. One implication of the proportional-hazard framework is that interactions between the covariates is implicit: the cross-partial of the hazard function with respect to two covariates is the hazard function times the product of the two coefficients on the covariates. For example, this multiplicative relationship between the covariates allows for price declines to have larger effects for riskier borrowers.

point estimates are to controlling for constant differences across regions.

Column 1 shows that the unadjusted differences in default rates across cohorts is even more pronounced when accounting for independent unobserved heterogeneity than the baseline results of column 1 of Table 3. I account for this by comparing the adjusted cohort coefficients in columns 2–5 to column 1 of Table 4. Including borrower and loan characteristics in column 2 explains 32% of cohort heterogeneity—the average decrease in the estimated cohort dummies. Controlling instead for 12-month price changes in column 3 reduces the residual difference in the default hazard across cohorts by an average of 68%. Conditioning on both price changes and loan and borrower characteristics in column 4 explains 92% of the cohort differentials in column 1. The total explanatory power of prices and observables is attenuated somewhat by including state fixed effects in column 5, where the combination of prices and observables explains 81% of the cohort pattern in column 1. Still, only the 2005 cohort is statistically significant at the 95% confidence level, and these standard errors are likely a lower bound because they do not allow for spatial correlations in default risk. Taking columns 4 and 5 together, as before, the 2005 cohort is the only borrower cohort to have a default hazard that is statistically distinguishable from the 2003 cohort hazard after adjusting for prices and loan and borrower observables, although these covariates explain 73% (column 5) to 81% (column 4) of the 2005 cohort coefficient in column 1. I conclude that the qualitative pattern of Table 3 is robust to allowing for independent unobserved heterogeneity: prices explain over 60% of cohort heterogeneity in default risk and combined with borrower and loan characteristics explain approximately 90% of the increase in defaults across cohorts.

1.5.2 Nonlinear Instrumental Variables Estimation

As discussed above, the interpretation of these results as causal requires the strong assumption that changes in the average default risk of a given area are not the cause of local price changes unless they are captured by loan and borrower covariates. Because defaults themselves cause price declines, this assumption is likely to be violated by any shock to area default risk. The demand shock resulting from the credit expansion may initially increase prices, and eventually a higher share of riskier borrowers may exacerbate price declines. In this way, if price changes are endogenous to subprime penetration and subprime growth reduces unobserved borrower quality, then the estimation would misattribute much of the increase in defaults to price changes instead of to differences in unobserved cohort quality. A second way that price declines may be endogenous to other factors that also affect default risk is from fluctuations in local labor market conditions. Adverse local labor shocks may simultaneously decrease prices (negative demand shock for owner-occupied housing) and increase defaults (negative income shock to existing mortgage borrowers).

The potential for changes in local house prices to themselves be a function of contemporaneous shocks to the default hazard through subprime lending or employment shocks necessitates instrumenting for prices. To instrument in this nonlinear setting, I use the control function approach (see Heckman and Robb, 1985). This estimator involves conditioning on a consistent estimate of the endogeneity in the endogenous explanatory variable and in a linear model is equivalent to two-stage least squares.[40]

To see why the control function approach solves the endogeneity problem, suppose again that there exists an omitted variable ω in the default hazard equation, which is not independent of X. Labeling the true hazard function $\tilde{\lambda}(\cdot)$, if

$$\tilde{\lambda}(X, \omega, t) = \exp(X\beta + \omega)\lambda_0(t) = e^{\omega}\exp(X\beta)\lambda_0(t).$$

If I do not control for ω in estimating this model, the resulting β coefficients will be estimating a different object than the marginal effect of X on the log hazard. Formally, the exogeneity condition introduced in equation (1.6) above now fails:

$$\begin{aligned} E\left[default_t - \lambda(X,t) \middle| X, \tau > t\right] &= E\left[\tilde{\lambda}(X,\omega,t) \middle| X\right] - \lambda(X,t) \\ &= E\left[e^{\omega} \middle| X\right]\lambda(X,t) - \lambda(X,t) \\ &= \left[\exp(X\beta + f(X)) - \exp(X\beta)\right]\lambda_0(t) \\ &\neq 0 \end{aligned}$$

where $E(e^{\omega}|X) \equiv f(X)$ because X and ω are not independent. Thus, under misspecification, the coefficients on X will not converge to the marginal effect of X on the log hazard and instead combine both the direct effect of X on default and the indirect effect of ω on default after projecting onto X.

Conditioning on an estimate of the endogenous component of X solves this problem. Let the

[40] Unlike a linear model, consistency of the control function approach in a nonlinear model relies on the instrument and the endogenous portion of the endogenous explanatory variable being independent (as opposed to just uncorrelated).

right-hand side endogenous variable be specified as

$$\Delta Prices = Z_1\Pi_1 + Z_2\Pi_2 + v$$

where the endogeneity problem arises because v and ω are not independent. The key identifying assumption is that the instruments Z_1 and included right-hand side controls Z_2 (the elements of X apart from $\Delta Prices$) are independent from v and ω. Conditioning on v then satisfies the exclusion restriction

$$\begin{aligned}
E\left[default_t - \lambda(X, v, t)\middle|X, v, \tau > t\right] &= E\left[\tilde{\lambda}(X, \omega, t)\middle|X, v\right] - \lambda(X, t, v) \\
&= E\left[e^{\omega}|v\right]\lambda(X, t, v) - \lambda(X, t, v) \\
&= \left(\exp(g(v)) - \exp(\rho_1 + \rho_2 v)\right)\exp(X\beta)\lambda_0(t)
\end{aligned}$$

where $g(v) \equiv E(e^{\omega}|v)$. If the conditional expectation $E(e^{\omega}|v) = \exp(\rho_1 + \rho_2 v)$, then this condition will hold, and controlling for a consistent estimate of v will be sufficient to allow estimation of the partial effect of X on the log hazard. This will be satisfied exactly if ω conditional on v is distributed normally: if $\omega|v \sim \mathcal{N}(\rho_2 v, 2\rho_1)$ then $e^{\omega}|v$ is distributed log normally with mean $E(e^{\omega}|v) = \exp(\rho_1 + \rho_2 v)$. If the conditional distribution of ω given v is non-normal, then controlling linearly for v in the hazard model relies on the quality of the linear model as a first-order approximation to the conditional mean function. As a robustness check, below I consider third- and fifth-order series approximations to the log of the conditional expectation function, e.g. $\log\left(E(e^{\omega}|v)\right) = \sum_{k=0}^{5}\rho_k v^k$ and find that the results are insensitive to this flexibility.

1.5.2.1 First Stage

The instrument set for the price change variable is the long-run cyclicality measure σ_g^P interacted with calendar-month indicator variables. The first stage for the 12-month price change is then

$$\Delta\log(HPI_{icgt}) = \sum_s \pi_s \sigma_g^P \cdot 1(s = t + t_0(i)) + Z'_{2,icgt}\pi_2 + v_{icgt} \tag{1.11}$$

where Z_{2icgt} contains the same covariates as equation (1.3) above—cohort effects, geographic fixed effects, loan and borrower characteristics, and the nonparametric baseline hazard function to ensure that predicted values from equation (1.11) are orthogonal to the other controls in equation (1.2).

The function $t_0(i)$ evaluates to the calendar time of loan i's origination date, and the π_s coefficients are turned on when the observation on loan i at t months after origination corresponds to calendar month s.

Table 5 reports the results from estimating equation (1.11) by OLS with standard errors clustered at the CBSA level. Column 1 includes just the instrument set and no other controls. The statistical relationship between actual price changes and the interactions between the cyclicality measure and calendar time is strong—the instruments explain 50% of the variation in twelve-month CBSA-level house price changes. Adding controls for the baseline hazard and CBSA fixed effects in column 2 improves the overall fit slightly (R^2 increases to .56). Including loan and borrower characteristics in column 3 does not affect the partial F-statistic, which tests the joint hypothesis that all of the coefficients on the instrument set are zero, suggesting that weak instruments are not a problem in this setting. The cohort coefficients in columns 2 and 3 illustrate that later cohorts were exposed to stronger price declines than earlier cohorts, in part by virtue of selection—younger loans are statistically more likely than older loans to not have terminated.

To provide intuition for how this instrument operates, I compute counterfactual price indices by regressing log home price indices on geographic fixed effects and an interaction of σ_g^P with calendar-month indicators as follows

$$\log(HPI_{gt}) = \alpha_g + \sum_s \pi_s \sigma_g^P \cdot 1(s = t) + u_{gt} \tag{1.12}$$

where HPI_{gt} is the value of the CoreLogic home price index in CBSA g in calendar month t. The estimated $\hat{\pi}_s$ shift the baseline log HPI of each CBSA (α_g) according to the cross-sectional relationship each calendar month between prices and 1980s price volatility.[41] Predicted values $\widehat{\log HPI}_{gt}$ from this regression provide an alternative time series of home prices in geography g based on the quasi-fixed tendency of home prices in geography g to cycle up and down.

Figure 6 shows the actual log home price series for 2003–2013 (left-hand panel) along with predicted values from equation (1.12) (right-hand panel). The left-hand panel shows that the actual HPI series are characterized by idiosyncratic deviations from the national trend, i.e. price shocks that potentially arise from such factors as local credit expansions and local labor market fluctuations

[41] It is worth pointing out that equation (1.12) does not control for main effects for each date. While this loads much of the national month-to-month variation in house prices onto the π_t, date effects are the very object the hazard model seeks to explain. As they aren't instruments and they don't belong in the second-stage, I purposefully omit them here.

that may also independently affect default rates. It is precisely the effects of these types of shocks that the instrument is designed to abstract from. Because nothing in equation (1.12) allows for differential price trends across CBSAs, the predicted time series in the right-hand panel all change in the same direction each month, differing only in the magnitude of the price change depending on their historical price volatility. To the extent that the actual price paths reflect time-varying local housing market changes, each line on the right is an estimate of the counterfactual price path that might occur absent local price shocks that are potentially driven by factors that also affect local default risk. Intuitively, my empirical strategy instruments for the actual price series on the left with the predicted price series on the right.

1.5.2.2 Exclusion Restriction

The necessary exclusion restriction for the IV results to be unbiased estimates of the causal effect of price changes is the independence of the size of a CBSA's 1980s price cycle (σ_g^P) from any other factors that affect default (besides prices). Note that with CBSA fixed effects, it is not a threat to identification if cyclical areas are fundamentally different from acyclical areas in some time-invariant way (e.g. an inherently risky area may always have both higher defaults and larger price swings). However, this exclusion restriction would be violated if pro-cyclical areas (high σ_g^P) have pro-cyclical trends in prices and the credit risk of borrowers. For example, if high-σ_g^P areas had more rapid subprime growth, then σ_g^P may proxy for changes in unobserved borrower quality in CBSA g. Similarly, if high-σ_g^P areas have greater unemployment rate fluctuations, these adverse shocks to local aggregate demand could increase defaults (through an income shock) and decrease prices (through a demand shock).

Figures 7 and 8 offer graphical evidence that subprime shares and unemployment rates—adjusting both for CBSA fixed effects—did not vary systematically with σ_g^P. The relevant period is different for each endogeneity concern. Figure 7 plots the annual adjusted subprime share of HMDA-covered mortgages originated in 2003–2007 by quartiles of σ_g^P. There is no apparent relationship between σ_g^P and subprime originations—places with historically large price cycles do not seem to have been any more prone to subprime credit expansion.[42] Figure 8 shows that the top quartile of σ_g^P had around a 1 percentage point *lower* unemployment rate in recession than the bottom quartile.

[42]Note that the same fact is not true about the relationship between subprime originations and the size of the late 2000s price cycle—areas that originated the highest share of subprime mortgages indeed had stronger (contemporaneous) price cycles, further evidence of the need for instrumenting.

Regression versions of these tests tell a somewhat different story. I test whether there is a first stage for the annual subprime share of all residential mortgage originations and the monthly unemployment rates by re-estimating equation (1.11), replacing the dependent variable with the subprime share of mortgages originated in each cohort and with the monthly unemployment rate in each CBSA. Consistent with Figure 7, the pattern of estimated coefficients $\hat{\pi}_t$ for the subprime share first stage is nearly completely flat and statistically insignificant, showing that loans from areas with higher historical price volatility were no more likely to have been originated during a relatively large subprime credit expansion. However, unlike in Figure 8, the estimated $\hat{\pi}_t$ in the unemployment regression mimic the national trends in the unemployment rate, with historically cyclical areas having differentially lower unemployment rates leading up to the recession and more quickly rising unemployment rates thereafter. This illustrates that areas with historically cyclical housing markets also have cyclical labor markets and that national labor market changes load onto the instruments. From this analysis, I conclude that σ_g^P successfully allows isolation of the effect of prices on defaults in a world where price declines and borrower quality are not jointly determined but that instrumental variables estimates are likely confounded by changing labor market conditions. An important caveat is that housing market changes can also affect labor markets (see Mian and Sufi, 2012). To the extent that the observed correlation between my instrument and labor market outcomes is an effect of the price cycle and not vice versa, then the instrument captures the total causal effect of price changes. However, because of the difficulty in ascertaining which caused which, I treat the relationship between the instrument and unemployment as a threat to validity. To account for this unemployment channel, I present additional control function specifications below that also control for the unemployment rate, thereby isolating the variation in prices that is not correlated with local labor market shocks or local subprime expansion.[43]

1.5.2.3 Control Function Results

Table 6 employs a nonlinear IV control function approach, which accounts for the endogeneity of price to the credit expansion by controlling for the first-stage residuals $\hat{v}_{icgt}$ in the default hazard

$$X'_{icgt}\beta = \gamma_c + W'_{B,i}\theta_B + W'_{L,i}\theta_L + \mu \cdot \Delta Prices_{icgt} + \kappa\hat{v}_{icgt} + \alpha_g \tag{1.13}$$

[43]The relationship between house price cyclicality and labor market cyclicality hints at the economics behind why some areas may be more cyclical than others. Areas with high housing supply elasticity, potentially arising from geographic constraints, land-use regulations, or credit market regulations, could be pro-cyclical in both markets. Similarly, areas with an industry mix that makes them particularly sensitive to recessions or commodity price shocks may experience coincident fluctuations in housing and labor.

where $\hat{v} = \Delta \log(HPI) - \Delta \widehat{\log HPI}$ and $\Delta \widehat{\log HPI}$ is fitted from equation (1.11). To account for the generated regressor problem in inference (Pagan, 1984 and Murphy and Topel, 1985), I also report bootstrapped standard errors in brackets below clustered standard errors. The generated regressor problem arises because v depends on an unknown parameter vector π, as seen in equation (1.11). Consistently estimating π in a first stage to generate $\hat{v}$ does not affect the consistency of parameters estimated in (1.13). However, by treating $\hat{v}$ as fixed, i.e failing to account for the correlation between the estimation error in $\hat{\pi}$ and the error in estimating β, the usual asymptotic standard errors will generally be understated unless $\kappa = 0$. The block bootstrap solves this by mimicking the data-generating process. In this setting, individual mortgages are resampled with replacement instead of month $\times$ loan observations being drawn with replacement as would be the approach of standard nonparametric bootstrap. The two stages (estimating $\hat{v}$ from (1.11) and estimating equation (1.13)) are then run on each bootstrapped sample and the resulting bootstrap standard errors are the empirical standard deviation of each element of β across 200 bootstrap replications.

Column 1 of Table 6 repeats column 6 of Table 3, controlling for the 12-month change in prices and not conditioning on borrower or loan observables W_B and W_L. Column 2 additionally controls for the residuals $\hat{v}_{icgt}$, estimated from OLS on equation (1.11) (omitting loan and borrower characteristics in the construction of the residuals). The coefficient on the price change variable is still large and significant—borrowers experiencing a 1% price decline over the previous year have a 4.4% higher conditional probability of default. The adjusted cohort differences are smaller in column 2 than column 1, meaning that after accounting for endogeneity, the role of prices in explaining the default pattern is larger. Comparing column 2 to the benchmark differences in cohort performance measured in column 1 of Table 3, controlling and instrumenting for prices without controlling for borrower or loan characteristics explains 60% of the difference in unadjusted cohort outcomes. The statistical significance of the coefficient κ on the residual is equivalent to a Hausman test for the endogeneity of price changes, similar to a Rivers and Vuong (1988) test for endogeneity in a probit model, confirming that price changes are endogenous.

To address the correlation between the instrument and local labor market shocks, column 3 also controls for the monthly CBSA unemployment rate, measured in percentage points.[44] Conditional on the covariates in the column 3 specification, a one percentage point increase in the local un-

[44]The sample size decreases slightly in specifications controlling for unemployment rate because of one CBSA for which BLS does not estimate monthly unemployment rates.

employment rate is associated with a decrease in the default hazard by 2%. The counterintuitive sign on the monthly unemployment rate and the increase in the magnitude of the coefficient on prices from –4.4 to 4.5 suggests that price results are not driven by correlation between price shocks and local labor shocks. The estimated differences in cohort quality in column 3 do not differ substantively from column 2.

The bootstrapped standard errors in columns 2 and 3 (in brackets) are in general much larger than the standard errors clustered at the CBSA level (in parentheses), representing a high degree of variability in the estimated residuals $\hat{v}$ across bootstrap samples using the control function approach. However, the relative stability of the coefficient magnitudes suggest that the patterns described above hold at least qualitatively. Further, because the asymptotic standard errors are correct under the null hypothesis $H_0 : \kappa = 0$, the conventional t-statistic on the fitted residuals is still a valid test of exogeneity.

The next three columns additionally control for borrower and loan characteristics. The estimated cohort effects in these specifications capture the latent quality of each cohort, i.e. the heterogeneity in cohort-level default rates not explained by ex-ante observable quality or price changes. Column 4 repeats column 7 of Table 3 for convenience, controlling for price changes in addition to all of the other controls. Column 5 reports control function estimates of this specification. The coefficient on the price change variable increases in magnitude from –3.9 to –4.6. The coefficient on the endogenous portion of the 12-month change in house prices is again positive and significant. Importantly, I cannot reject that each of the cohort latent quality measures is statistically indistinguishable from zero with the exception of the 2005 cohort, as before. Moreover, the estimated cohort effects in column 4 are each smaller than those in column 3 which treat prices as exogenous. Column 6 again controls for the monthly unemployment rate. The magnitude of unemployment on default is almost identical as in column 3, suggesting that the unemployment rate does not interact meaningfully with loan and borrower characteristics. The price effects—both the main effect and the residuals—are strengthened by the inclusion of the unemployment rate control, although this difference is not statistically significant. Each of the cohort effects is attenuated slightly from column 5. The specifications in columns 5 and 6 both explain 95% of the unadjusted differences in cohort default rates in column 1 of Table 3.

Interestingly, the bootstrapped standard errors in columns 5 and 6 are much more similar to their asymptotic counterparts than the bootstrapped standard errors of columns 2 and 3. Unlike columns 2 and 3, the results of columns 5 and 6 are robust to bootstrapping the standard errors. This suggests that much of the instability of the bootstrap estimates in columns 2 and 3 is driven by not controlling for loan and borrower characteristics, which explain a substantial amount of individual heterogeneity in default risk.

A consistent pattern in Table 6 is that instrumenting (columns 2, 3, 5, and 6) increases the magnitude of the estimated effect of price changes relative to not instrumenting (columns 1 and 4). An explanation for this is the positive sign on the estimated coefficient $\hat{\kappa}$ on the residuals. While the partial effect of a price shock v on the log of the default hazard—equal to $\mu + \kappa$ because the residuals enter into the $X\beta$ index through both $\Delta Prices$ and v—is strongly negative in each specification, the effect of an exogenous change in prices captured by μ alone is much greater. This is consistent with some degree of treatment effect heterogeneity—if price declines arising from shocks that are correlated with default risk (e.g. credit market changes) have a weaker effect on defaults than price declines induced by, for example, national price declines unrelated to local credit market fluctuations, then isolating the exogenous variation in prices would increase the estimated price effect.

Table 7 addresses the possibility that the conditional distribution of the endogeneity is misspecified. As mentioned above, controlling for v linearly in $X\beta$ relies on the assumption that the omitted default risk factors ω are distributed normally condition on v. In general, if $\omega|v \not\sim \mathcal{N}$ then $E(e^{\omega}|v) = g(v) \neq \rho_1 + \rho_2 v$. In this case, the specification of $X\beta$ needs to be augmented to include a consistent estimate of $\log(g(v))$, which I approximate using third- and fifth-order polynomials in the fitted residuals, e.g. $\log\left(g(v_{icgt})\right) = \sum_{k=0}^{5} \rho_k \hat{v}_{icgt}^{k}$. Columns 1–3 do not control for borrower or loan characteristics. Column 1 is repeated from column 1 of Table 6 for convenience. Column 2 adds the residuals squared and the residuals cubed. These coefficients are strongly significant, and a likelihood ratio test for the hypothesis that $\rho_2 = \rho_3 = 0$ rejects, pointing to likely non-normality of the unobserved heterogeneity that is correlated with price shocks. However, the slope coefficients are relatively unaffected from the additional flexibility in the estimate of $\log\left(E(e^{\omega}|v)\right)$. Column 3 adds fourth- and fifth-order terms, which again do not noticeably affect the estimated effect of prices or differences in the latent quality of cohorts. The estimated coefficients $\hat{\rho}$ on the powers of the residuals in column 3 are very imprecise, and a likelihood ratio test fails to reject that $\rho_4 = \rho_5 = 0$. Columns 4–6 in Table 7 repeat the specifications in columns 1–3, additionally controlling for borrower and loan characteristics. The same patterns are apparent: powers of the residuals are jointly significant, rejecting the exogeneity of price changes, and the estimated effects of the covariates are relatively unchanged.

The control function results of Tables 5 and 6 are consistent with the results of Table 3, providing evidence that there is a large causal effect of price declines on defaults. Even after accounting for the endogeneity of the effect of prices on default risk and controlling for local labor market conditions, there is little evidence that unobserved borrower quality declined across 2003–2007 cohorts. Comparing the asymptotic and bootstrapped standard errors, that pattern of Table 6 holds that the bootstrapped standard errors are greatly affected by the inclusion of micro-level covariates as controls. The bootstrapped standard errors in columns 1–3 are often an order of magnitude larger than the corresponding asymptotic ones, while the bootstrapped standard errors of columns 4–6 are an average of only 29% higher than the asymptotic standard errors.

In summary, this section was concerned with determining how much of the pattern across origination cohorts in default rates was due to differences in the observed characteristics of mortgage borrowers in each cohort—both the creditworthiness of the individual borrowers and the characteristics of their mortgages—and differences in their exposure to price declines. The results confirm that prices and mortgage characteristics are both are important, with price changes causally explaining at least 60% of the increase in cohort default rates.

1.6 Mechanisms

I now turn towards identifying the causal mechanisms through which prices affect default rates by testing for negative equity having a causal impact on defaults and whether this explains cohort heterogeneity. The intuition offered above centers around the differential effect of price declines on later cohorts in pushing them underwater, as seen in Figure 2. Mortgage borrowers who are underwater have elevated default risk. Distressed borrowers (i.e. borrowers unable to make their monthly mortgage payments) who have positive equity have two main alternatives to default. First, if interest rates have gone down or if borrowers qualify for a lower interest rate because they have more equity from paid down mortgage principal and accumulated price appreciation, they can refinance into a mortgage with a lower monthly payment, using the new mortgage to repay the original one.[45] Second, they can sell their home and use the proceeds to pay off their outstanding

[45]Relatedly, a home equity line of credit can also be used to borrow additional funds secured by unrealized capital gains. These funds can be used to temporarily make mortgage payments.

mortgage debt and move into a more affordable housing situation. Neither of these options is readily available to distressed borrowers who are underwater. Lenders are normally unwilling to originate a refinance mortgage to someone who has zero equity. let alone negative equity. Selling a house secured by a mortgage in a negative equity position (known as a short sale) requires either coming up with sufficient cash to pay the shortfall between the sale price and the outstanding debt or working with the lender to secure forgiveness of the remaining debt. By definition, distressed borrowers are unlikely to have ample savings, making the former unlikely. Lenders are also wary of agreeing to short sales, partly because of asymmetric information about the borrower's current and future finances. An additional source of elevated default risk comes from the possibility that underwater borrowers will default strategically.[46]

Empirically testing that the reason price declines explain the bulk of cohort heterogeneity is through the prevalence of negative equity presents several practical challenges. First, the extent to which borrowers are current with their monthly payments is related to their unobserved quality. I instrument for the actual balance of the mortgage with the scheduled balance calculated using the origination interest rate as if the borrower had paid back a 30-year fixed-rate mortgage on schedule. Second, constructing a measure of negative equity status requires knowing the current market value of the home, an unknown (and endogenous) quantity that must be estimated by the borrower as well as the econometrician. CoreLogic provides such a measure using their Automated Valuation Model that imputes property values in each month for each subprime mortgage in the data. As this estimated value is partly a function of nearby market prices and therefore affected by CBSA-level shocks, I instrument for this valuation using the origination loan amount and counterfactual price indices computed using the historical volatility instruments. Third, because the prepayment obstacles faced by borrowers depend on the total debt of all loans secured against their home, measuring negative equity necessitates knowing updated information about additional liens. Not observing updated information on the outstanding balance of additional liens, I assume that all second mortgages have not been paid down. Although this introduces additional measurement error into the estimated balances, which are already affected by local public records access policies, instrumenting for outstanding balances using scheduled balances solves this problem.

[46]Strategic default is when a borrower who has available cash flow to make mortgage payments defaults anyway, exercising a put option on the property. This is optimal if the option value of holding onto the property (i.e. expected future price appreciation) is lower than value of discharging the debt, net any cost of defaulting (see Foote et al., 2008). In other words, borrowers may find it advantageous to default if they do not expect future prices to rise quickly enough.

I define the variable $Underwater_{it}$ for whether the current CLTV of a loan, estimated by CoreLogic based on the outstanding debt owed to all outstanding liens and contemporaneous market conditions, is greater than 100%. I first present estimates that do not account for the endogeneity of CLTV. Table 8 contains default hazard specifications of the form above, replacing $\Delta Prices$ with functions of CLTV

$$X_{icgt}'\beta = \gamma_c + W_{B,i}'\theta_B + W_{L,i}'\theta_L + \eta' q(CLTV_{it}) + \alpha_g. \tag{1.14}$$

Controlling for $q(CLTV) = 1(CLTV > 1) = Underwater$ in addition to loan and borrower characteristics and CBSA fixed effects in column 1 of Table 8 shows that underwater mortgages have more than double the conditional default probability of mortgages that are not underwater. There is substantial unexplained cohort heterogeneity in column 1—even after adjusting for location, mortgage age, borrower and loan characteristics, and the estimated negative equity status, differences in cohort default rates relative to the 2003 cohort are all positive and significant, with the exception of the 2007 cohort. Compared with column 7 of Table 3, the underwater indicator variable explains much less cohort heterogeneity than the 12-month change in prices. Columns 2 of Table 8 tests whether this is driven by the functional form restriction on $q(\cdot)$ by controlling for a linear spline in the current CLTV that allows for a location and scale shift in the effect of $CLTV$ in several bins:

$$q(CLTV_{it}) = \sum_{j=1}^{J} 1(CLTV_{it} \in C_j) \times (a_j + b_j CLTV_{it}) \tag{1.15}$$

where j indexes the set C consisting of J CLTV intervals $\{[0, 80), [80, 85), [85, 90), ..., [150, \infty)\}$. Adding flexibility in the specification of the leverage function $q(\cdot)$ further decreases the adjusted differences across cohorts but only explains on average an additional 8% of the differences in the latent quality of cohorts. The specification in column 2 explains 62% of the cohort-level differences in default.

To see whether prices still explain cohort heterogeneity even conditional on underwater, i.e. to test whether the effect of prices is driven entirely by negative equity, column 3 additionally controls for the twelve-month change in log HPI. Adding in the price change variable in addition to the linear spline controls significantly affects the estimated cohort heterogeneity relative to column 2 but also relative to column 7 of Table 3, which is identical to column 3 except for the inclusion of $q(CLTV)$. This suggests that CLTVs and prices interact in explaining defaults. Controlling for both price changes and current CLTVs reduces the 2006 and 2007 cohort differences to be strongly negative—controlling for a flexible function of their relative equity, the price changes they faced, and loan and borrower characteristics, 2006–2007 borrowers defaulted less than would be expected. The estimated latent quality of the 2004–2005 cohorts positive and significant, with the 2005 estimate smaller and the 2004 result larger than the results in Table 3 that do not control for current CLTVs. The coefficient on the price change variable is large and significant.

There is a large relationship between defaults and negative equity and evidence that prices also affect defaults in other ways than through negative equity. Still, caution is required interpreting these results because mark-to-market leverage (CLTV) could be correlated with unobserved borrower quality. I now discuss an instrumental-variables strategy to account for this endogeneity.

1.6.1 Instrumenting for Loan-to-Value Ratios

The main obstacle in interpreting the results in columns 1–3 of Table 8 is the endogeneity of CLTVs, which are the ratio of loan principal balances and property values. To the extent that borrowers whose unobserved quality is low (high) pay back their mortgages more slowly (rapidly), loan balances (and hence CLTVs) will be determined in part by unobserved borrower quality. Similarly, borrowers with lower unobserved quality may take out mortgages with slow amortization schedules that leave them more likely to be underwater. To address the endogeneity of CLTV numerators, I calculate the scheduled loan principal amount at each month since origination if borrower had taken out a 30-year fixed interest rate loan with same origination interest rate and purchase price and was current on all payments time.[47] Using the amortization formula,

$$Scheduled\ Principal_{it} = M_i \left((1+r_i)^t - \frac{(1+r_i)^{360}\left((1+r_i)^t - 1\right)}{(1+r_i)^{360} - 1} \right)$$

where M_i is the purchase price of property i, t is the loan age in months, and r_i is the origination interest rate divided by 12.

To account for the endogeneity in purchase prices, I compute what the purchase price of the home would have been if the borrower had taken out only a first-mortgage for the same dollar amount at the conforming loan limit (80% of purchase price). In logs, using this predicted purchase price is equivalent to using the log of the origination amount as an instrument. Finally, with the predicted

[47] Cunningham and Reed (2013) refer to this as a synthetic mortgage IV strategy.

home price indices, I can calculate an alternative measure of the change in a property's value since origination using the predicted HPI series

$$\widehat{Appreciation}_{igt} = \widehat{\log HPI}_{igt} - \widehat{\log HPI}_{ig1}$$

where $\widehat{\log HPI}$ are the predicted values from estimating equation (1.12).

Before presenting first stage results, in FIgure 9 I illustrate graphically the statistical relationship between each of the three instruments and the corresponding component of CLTVs. Diagonal lines depict the fitted bivariate linear regression line. Panel I plots actual log principal balances versus scheduled log principal balances. The fit is very strong and the slope of the bivariate regression line is close to 1, showing the tight relationship between traditional amortization schedules and loan balances. The most noticeable deviation is the presence of many outliers well below the regression line, representing people that paid their mortgages back faster than scheduled. Instrumenting will address the possibility that their faster payback is a signal of these borrowers' unobserved (high) quality. Panel II plots actual log sale prices against log origination amounts. The average relationship between origination balances and actual sale prices is not far off from a setting where all borrowers took out mortgages at 80% of the sale price of the home, in which case there would be a perfect fit between log origination amount and log sale price with an intercept of log(1.25) and a slope of 1. The most obvious outliers are those well above the regression line—borrowers who took out mortgages with much lower leverage (i.e. through a larger downpayment in the case of sales or from accumulated equity in the case of refinances). Using log origination amounts as an instrument to explain CLTVs will account for any correlation between actual sale prices, initial leverage, and unobserved borrower quality. Panel III plots assessed property values against counterfactual property values

$$\widehat{Value} = 1.25 \times Origination\ Amount \times \exp(\widehat{Appreciation})$$

to show the predictive power of the generated instrument $\widehat{Appreciation}$. The workhorse behind this relationship is the long-run price cyclicality instrument σ_g^P used to predict HPI values and subsequently impute appreciation-since-origination and corresponding counterfactual property values. There is a clear positive relationship between counterfactual property values and assessed values. Positive deviations from the regression line represent homes in areas and months with much higher prices than would be predicted based on the 1980s price cycle of that city. Negative deviations represent homes where price declines have been more acute than expected a priori. Instrumenting for actual assessed values will address the potential for these price changes to be correlated with unobserved borrower default risk.

The first stage for CLTV is a linear regression of CLTV on the scheduled loan balance, the loan origination amount, predicted appreciation using the counterfactual price series, and the usual controls Z_2

$$CLTV_{icgt} = Z'_{1,igt}\Upsilon_1 + Z'_{2,icgt}\Upsilon_2 + \nu_{icgt} \tag{1.16}$$

where the instrument set consists of

$$Z_{1,igt} = \Big(\log(Scheduled\, Principal_{it}) \quad \log(Origination\, Amount_i) \quad \widehat{Appreciation}_{igt} \Big).$$

Table 9 reports the results of estimating equation (1.16) by OLS with clustered standard errors. Note that missing data—loans for which CoreLogic has not estimated a contemporaneous CLTV in a given month—reduces the sample size of specifications involving CLTV from 1.2 to 1.0 million monthly loan observations. Column 1 reports results of regressing $CLTV$ on Z_1 without controlling for Z_2. The relationship between each of the instruments and CLTV values is large and very precisely estimated. Mortgages with higher origination amounts (positive predictors of sale prices) have lower CLTVs. Mortgages with higher scheduled principal balances have higher CLTVs. Mortgages with higher predicted appreciation have lower CLTVs. Adding cohort indicator variables, baseline hazard controls, and CBSA fixed effects in column 2 strengthens the estimated effect of origination amounts and scheduled principal and attenuates the effect of predicted appreciation on the CoreLogic contemporaneous CLTVs. The cohort pattern confirms the trends in median CLTVs plotted in Figure 2, showing that later cohorts have much higher CLTVs. Successively controlling for borrower and loan characteristics in column 3 and price changes in column 4 continues the trend. The instruments are still powerful predictors of CLTVs. Column 5 additionally controls for the monthly CBSA unemployment rate. Local labor market conditions are clearly correlated with CLTVs: the coefficient on the unemployment rate suggests that the equity share of property values in areas with high unemployment rate is lower. Controlling for the unemployment rate, the predicted appreciation instrument is no longer significant. Still, the partial F-statistic for the joint significance of the instruments is above 200 in every column.

Columns 4–6 of Table 8 report the results of estimating the default hazard function after incorporating $\hat{\nu}_{icgt}$ from equation (1.16) into the linear index $X\beta$ in equation (1.14).[48] Column 4 includes the underwater indicator variable as a parsimonious summary of the causal influence of negative equity on default conditional on the CLTV residuals, price changes, and price change residuals. Columns 5 and 6 instead control for a linear spline in q. The estimated effect of prices is large and significant across all specifications, showing an elasticity of default with respect to price declines of -3 to -5, meaning that for a fixed CLTV, a 1% price decline increases the default hazard by 3–5%. The effect of being underwater on default is still significant but greatly attenuated from column 1, suggesting that holding prices fixed, a mortgage being underwater *causes* the default hazard to be 33% higher (28 log points) than that of above-water mortgages. This suggests that some of the performance differences across cohorts that columns 1–3 attributed to negative equity were actually unobserved differences in borrower quality across cohorts that affected both defaults and equity. Indeed, the CLTV residuals are significant in columns 4–6, rejecting the null hypothesis that CLTVs are exogenous.

Comparing columns 3 and 5, the estimated cohort differences after controlling and instrumenting for price changes and mark-to-market leverage are slightly smaller than the corresponding estimates column 3 that do not account for the endogeneity of prices or CLTVs. Continuing a trend in my findings, the specification in column 5 is more successful at explaining the default rates of later cohorts than earlier cohorts, suggesting that that negative equity was a more important factor in late-cohort defaults than early cohort defaults. While highly predictive of individual defaults, the smaller effect of CLTV controls on earlier cohort default rates is consistent with earlier cohorts' CLTVs not having increased as much (see Figure 2).

Because local labor market fluctuations are not excludable from my instrument, I control directly for the unemployment rate in column 6 of Table 8. As in Table 6, conditional on all of the other controls, mortgages in cities with increased unemployment rates are slightly less likely to default—a one percentage point increase in the local unemployment rate decreases the default hazard by 5%. Accounting for local labor market fluctuations does not materially affect the estimated coefficients on prices or CLTV residuals. However, including the unemployment rate decreases the measure of the difference in latent quality between the 2003 and 2004–2005 cohorts enough to be statistically

[48] Imbens and Wooldridge (2007) discuss the control function approach when the estimating equation contains several non-linear functions of the right-hand side endogenous variable. Under the assumption that the unobserved component of default risk is independent of the instruments (the control function exclusion restriction), controlling for the fitted residuals of CLTV is sufficient to instrument for any function of CLTV.

insignificant.

Taken together, the results of Table 8 provide several explanations about the mechanisms through which price declines cause defaults, decomposing cohort-level differences in default rates into four factors: borrower and loan characteristics, price declines, and local economic conditions. Negative equity is a prominent channel and explains much of the the relationship between cohort default rates and price changes, especially among later borrower cohorts. Another important factor is unemployment, which may cause and be caused by price declines (Mian and Sufi, 2012). Nevertheless, prices affect default risk in other ways besides their effect on default through equity and their correlation with local economic conditions. The role of price expectations is one likely explanation for prices having such a strong relationship with default even conditional on negative equity. If buyers' expectations of future prices are correlated with recent price changes, even above-water borrowers wishing to sell in areas experiencing recent price declines may be unable to in the face of a thin market of patient buyers. Still, there is strong evidence that negative equity is responsible for much of the effect of prices on defaults and that the differential prevalence of negative equity across cohorts explains a significant portion of the observed increase in cohort-level default rates.

1.7 Estimating Counterfactual Default Rates

Using the control function specification estimated in column 5 of Table 6 as my preferred specification, I calculate average default rates for each cohort using counterfactual explanatory variables as an estimate of the impact of the price and mortgage characteristics channels. Using the estimated coefficients, predicted values $\hat{h}$ are an estimate of the probability each loan defaulted for each month it existed. By equation (1.4),

$$\hat{h}_{icgt} = 1 - \exp(-\exp(X'_{icgt}\hat{\beta} + \hat{\psi}_t)$$

where $\hat{\psi}_t$ are nonparametric estimates of the log baseline hazard function between time $t-1$ and t as discussed in Section 1.4.1. I aggregate these individual default probabilities to calculate cohort-level average default rates, which I annualize multiplying by twelve. The predicted average annual default rate for cohort c is then defined as

$$\widehat{\text{Default Rate}}_c = \frac{12}{N_c} \times \sum_{t \leq 60} \hat{h}_{icgt}$$

where N_c is the number of cohort-c monthly loan observations in the sample of loans within five years of origination. I limit the sample to observations on loans within five years of origination to facilitate comparisons across cohorts. Because I define default to occur the first month that a mortgage is marked as in foreclosure or real-estate owned, this rate is similar to the average number of foreclosure starts in each month divided by the number of loans that were extant during that month.

Table 10 shows the counterfactual default rates for eight scenarios, each representing a different combination of counterfactual price paths and loan characteristics. The first row reports the actual default rates for each cohort. The actual spread between the default rates of 2003 and 2006 mortgages was 8.2 percentage points. The model's predicted default rates using observed covariates (not shown) match the actual default rates to 2–3 decimal places, suggesting that this parsimonious model fits the data quite well. Rows 2 and 3 estimate the average default rates that would have prevailed if all mortgages had the characteristics reported in Table 2 of the average 2003 (row 2) or 2006 (row 3) mortgage. Default rates would have been lower if the characteristics of mortgages had not changed over time, especially for later borrower cohorts. If all borrowers had taken out the average 2006 mortgage, row 3 shows that default rates would have been roughly one percentage point higher for 2003–2005 cohorts and lower for 2006–2007 cohorts. The spread between the 2003 and 2006 cohorts is cut in half by fixing mortgage characteristics. However, even if the composition of mortgage products did not change from 2003–2006—a conceptual upper bound on the effect of stricter mortgage regulation, the 2006 cohort would have still defaulted 3.7 percentage points more frequently than the 2003 cohort.[49]

The remaining rows experiment with counterfactual price paths. Rows 4–6 use actual individual loan characteristics and three alternative price scenarios. Row 4 assigns each loan to have the average price change that 2003-cohort loans faced at the same number of months since origination. Row 5 does the same exercise using the prices to which 2006-cohort loans were exposed, and row 6 looks at the effect of flat prices—0% price growth over the life of the mortgage. As expected, mortgages from every cohort would have defaulted much less if they had experienced several years of rapid price appreciation, as did 2003-cohort mortgages. Ceteris paribus, if 2006-cohort mortgages had faced the same prices that the average 2003 mortgage did, their default rate would have been 5.6% instead of 12%. Similarly, if 2003-cohort mortgages had faced the prices that the average

[49]Note that this statement assumes that holding mortgage product characteristics fixed would not have affected aggregate prices.

2006 mortgage faced, their default rate would have been 8.5% instead of 4.2%. The counterfactual default rates for the scenario in which there is no price growth is predictably in between the 2003 and 2006 price scenarios. The spread between the 2003 and 2006 cohorts seen in row 1 is mostly gone in rows 4–6, showing that if they had faced the same prices, the 2006 cohort default rate would have been at most 2.5 percentage points higher than the 2003 default rate.

The final two rows report default rates under the counterfactual of constant prices and mortgage characteristics. The combination of fixed prices and mortgage characteristics explains the entire difference in the unadjusted cohort default rates of column 1, with the 2006 cohort predicted to *outperform* the 2003 cohort if both had faced the same (zero) price growth and had taken out mortgages with the characteristics of either the average 2003-cohort mortgage (row 7) or the average 2006-cohort mortgage (row 8). As a measure of the latent quality of each of these cohorts, rows 7 and 8 suggest that there were no important declines in unobserved borrower quality across subprime cohorts. Using the zero price growth scenarios as a benchmark, it seems that the low and high actual default rates experienced by the 2003 and 2006 cohorts, respectively, were not particularly representative of the relative quality of these cohorts. Intuitively, this makes sense—by historical standards, neither the 2003 nor 2006 price paths seem to have been particularly normal.

1.8 Conclusion

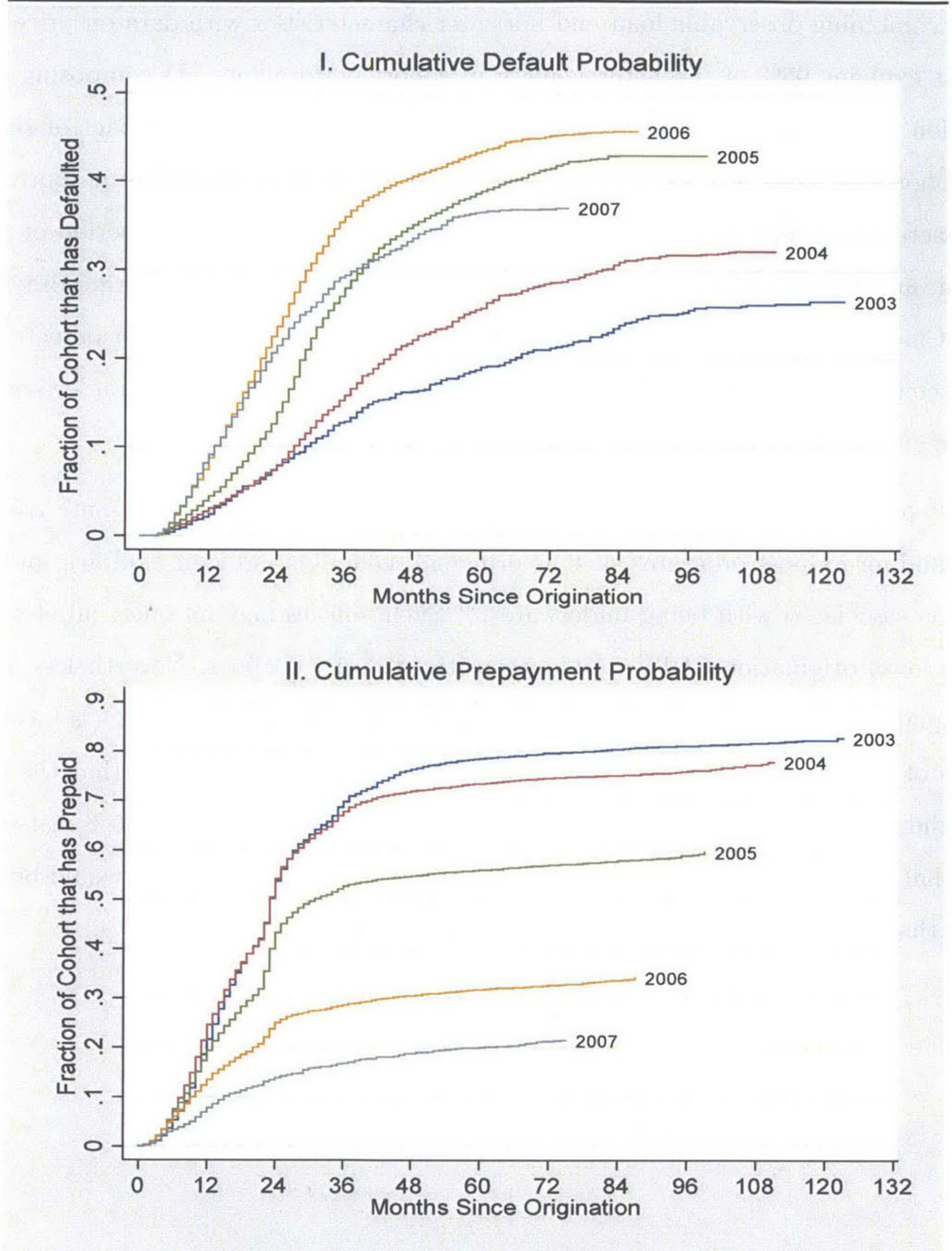
There has been an active debate about the surge in the subprime default rate in the mid- to late-2000s, with blame being placed on risky mortgage products, risky borrowers, and price declines. The accompanying analysis has focused on contrasting the relative performance of late and early cohorts to tease out these stories. Diverse views of the cause of this deterioration in cohort-level mortgage outcomes have motivated strong opinions about the appropriate regulator response to the subprime default crisis. Advocates of stricter mortgage lending regulation argue that the cohort pattern represents a deterioration in underwriting standards over time, i.e. the lending of riskier mortgage products to riskier borrowers, and that these looser standards were the main precipitating factor in the crash.

This paper demonstrates why the cohort comparison is potentially misleading: cohorts may differ not only in their composition (loan and borrower characteristics) but crucially in the degree to which they were affected by price fluctuations. I ascertain the relative contribution of each of these factors by combining observable loan and borrower characteristics with data on price changes in a model that explains 95% of the heterogeneity in cohort performance. Decomposing the observed deterioration in subprime loan performance, I find that the differential impact of the price cycle on later cohorts explains 60% of the rapid rise in default rates across subprime borrower cohorts. Loan characteristics, especially whether the mortgage had an interest-only period or was not fully amortizing, are important as well and explain 30% of the observed default rate differences across cohorts. Changing borrower characteristics, on the other hand, had little detectable effect on cohort outcomes. While quite predictive of individual default, borrower characteristics simply did not change enough across cohorts to explain the increase in defaults.

The results of this paper suggest a scope for underwriting standards to ex-ante affect mortgage outcomes and for ex-post programs such as principal reduction and loan modifications that reduce the frictions associated with being underwater.[50] Such policies may interact: all else equal, mortgages with lower origination CLTVs are less sensitive to price declines. Nevertheless, the view that borrower quality declined across subprime cohorts on unobservable dimensions is inconsistent with the results of this paper. I find that if 2006 borrowers had faced the prices that the average 2003 borrower did, 2006 borrowers would have had an annual default rate of 5.6% instead of 12%. I conclude that a 2003 borrower taking out the average 2006 mortgage in 2006 would be no less likely to default than a 2006 borrower in the same circumstances.

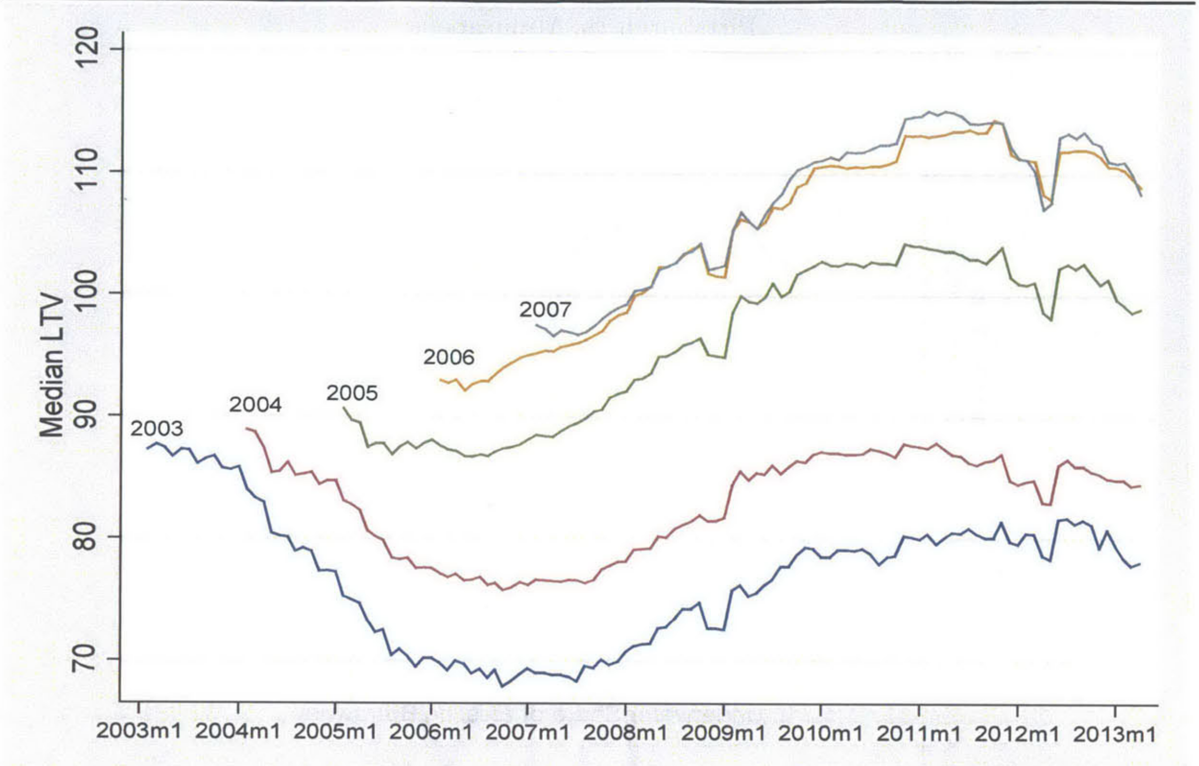
[50] Of course, the presence of ex-post remedies may induce moral hazard. See Mayer and Hubbard (2009), Wheaton (2010), Feldstein (2011), and the enacted Home Affordable Modification Program for examples of loan modification programs and proposals, many designed to preserve incentives for responsible borrowing and maintenance.

Figure 1. Cumulative Default Probability by Origination Cohort



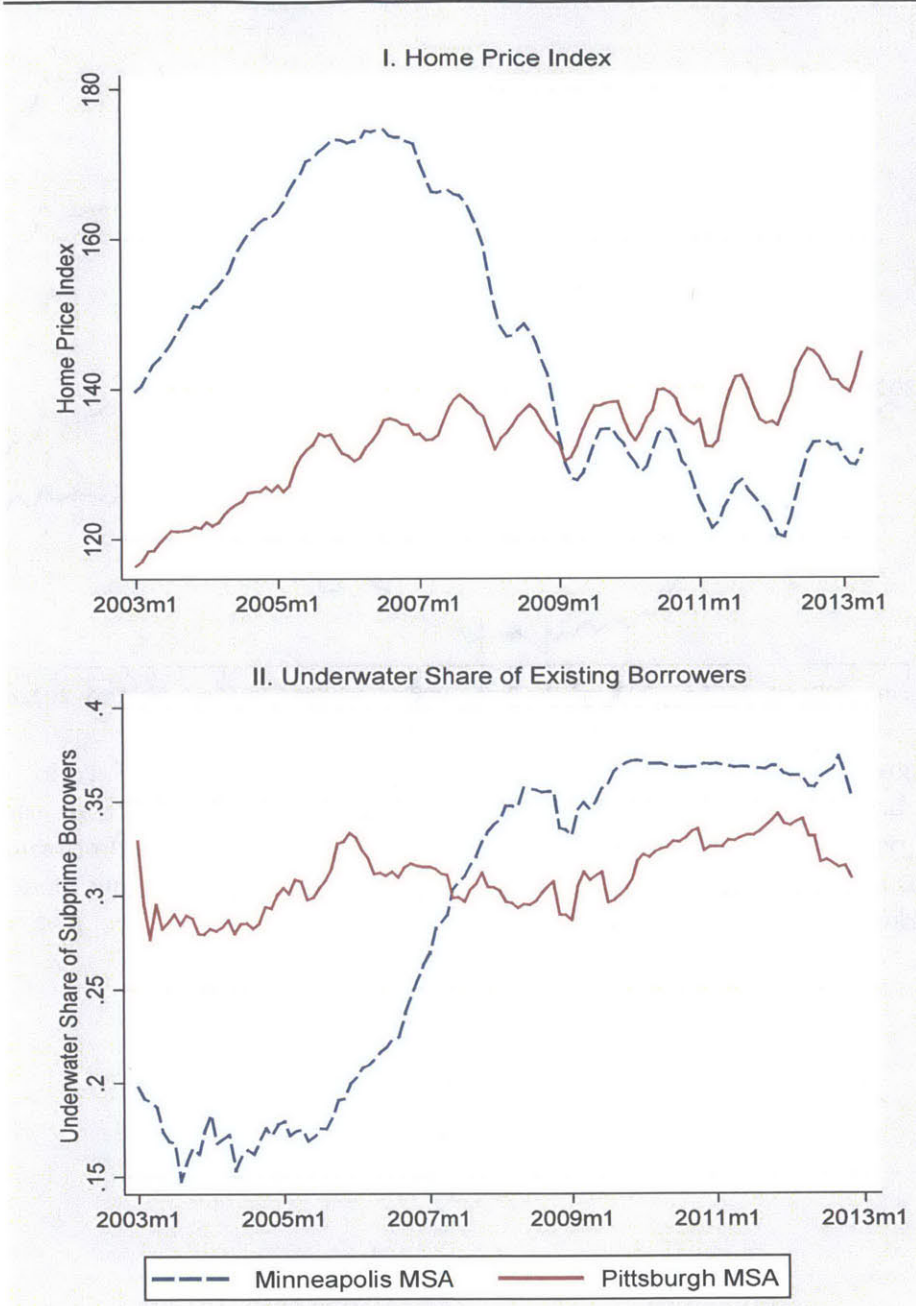
Notes: Figure plots the fraction of each cohort that has terminated by default (left panel) or prepayment (right panel) within a given number of months since origination. Default is measured as the first time that a loan's delinquency status is marked as in foreclosure or real-estate owned provided it ultimately terminated without being paid off in full. Prepayment means repayment in full, i.e. through refinancing or selling.

Figure 2. Median Combined Loan-to-Value Ratio Over Time by Cohort



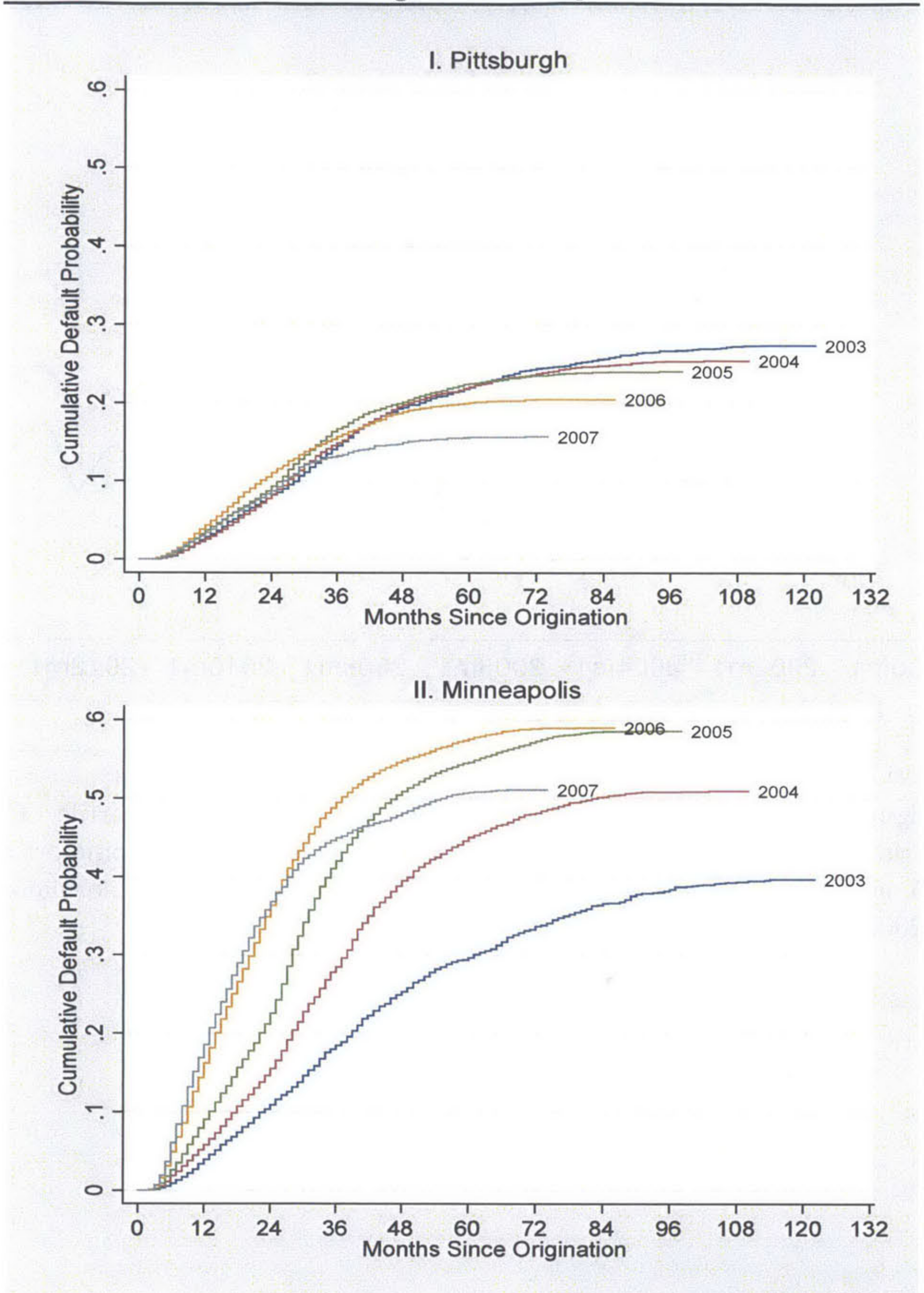
Notes: Figure shows the median current combined loan-to-value ratio (CLTV) of subprime borrowers for existing subprime mortgages in each cohort in each calendar month in percentage points. Current CLTVs are calculated by LoanPerformance as the total outstanding principal on a loan divided by an automated assessing model's estimate of the market value of each home.

Figure 3. Prices and Negative Equity Prevalence:
Pittsburgh vs. Minneapolis



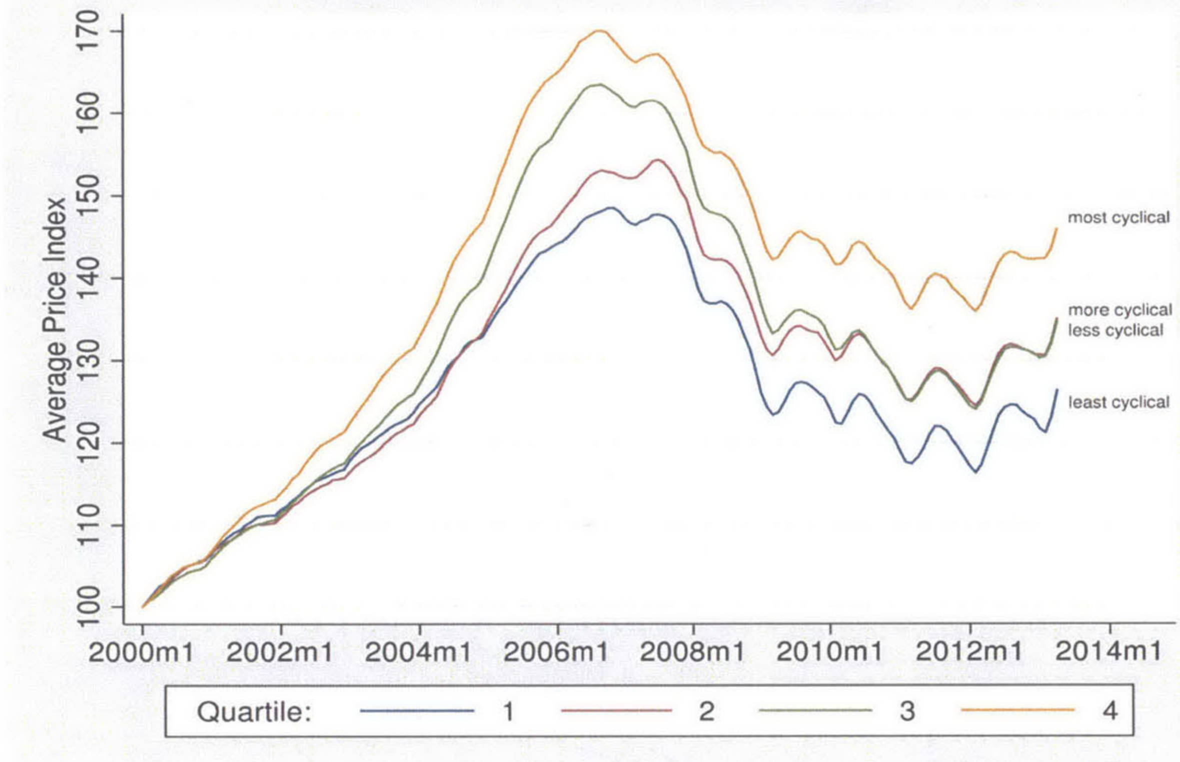
Notes: Top panel shows the CoreLogic repeat-sales home price index for the Minneapolis and Pittsburgh Metropolitan Statistical Areas (MSA). Both series have been normalized to 100 in January 2000. Bottom panel shows shows the fraction of all outstanding subprime borrowers that were underwater in each calendar month in the indicated MSA. Underwater is determined by the current combined loan-to-value ratio (CLTV) for a loan being above 100%. CLTVs are calculated by CoreLogic as the total outstanding principal on a loan divided by an automated assessing model's estimate of the value of each home.

Figure 4. Cumulative Default Probabilities:
Pittsburgh vs. Minneapolis



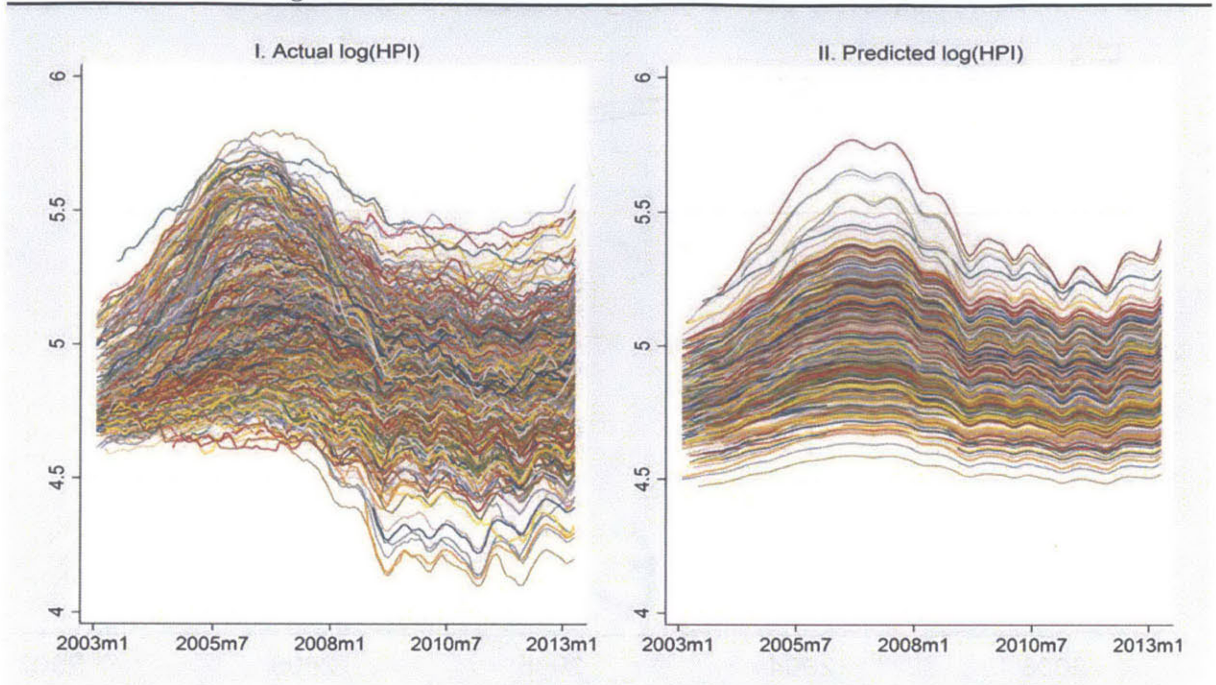
Notes: Graphs show the cumualtive default probability of each cohort in the Pittsburgh and Minneapolis CBSAs, respectively. Each line shows the fraction of that cohort that had defaulted within a given number of months since origination. See Figure 1 notes for more details.

Figure 5. Persistence of House Price Cyclicality:
Average Home Price Index by Quartile of σ^P



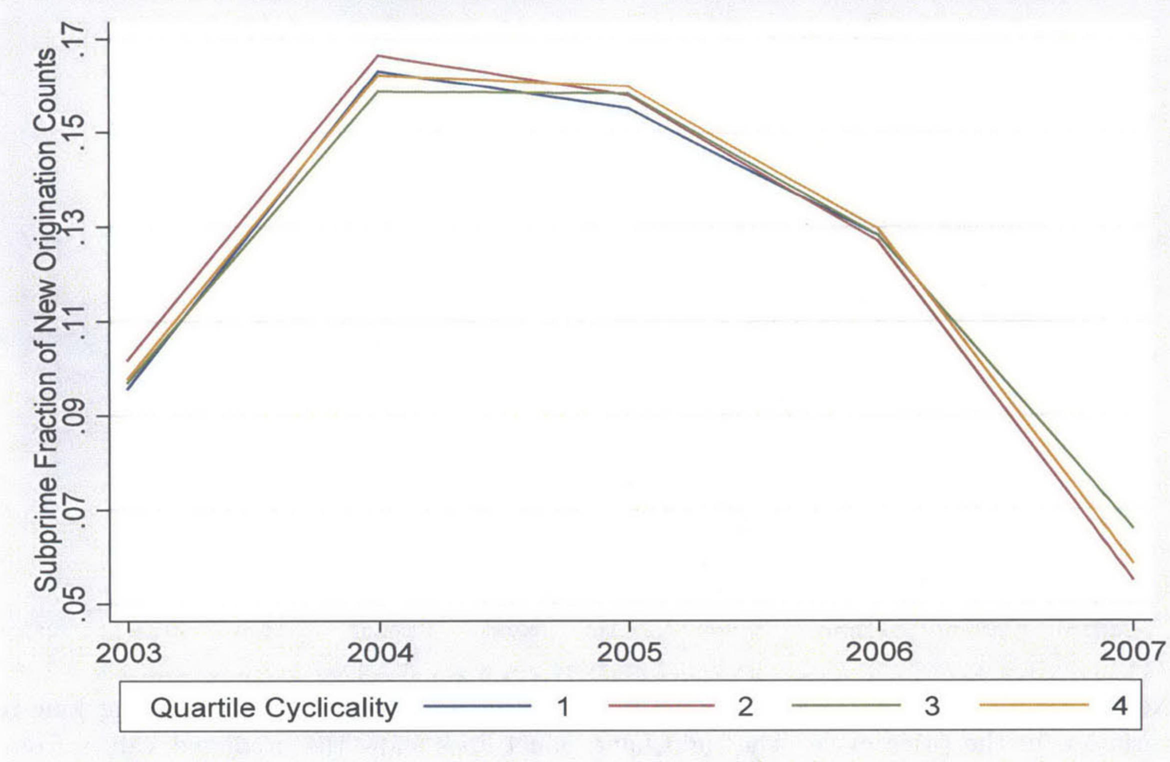
Notes: Figure plots month average HPI values by cyclicality quartile. Cyclicality is measured as the standard deviation of one month changes to the log home price index from 1980-1995, as defined in equation (5) in the text. Each series has been normalized to 100 in January 2000.

Figure 6. Observed and Predicted Home Price Indices



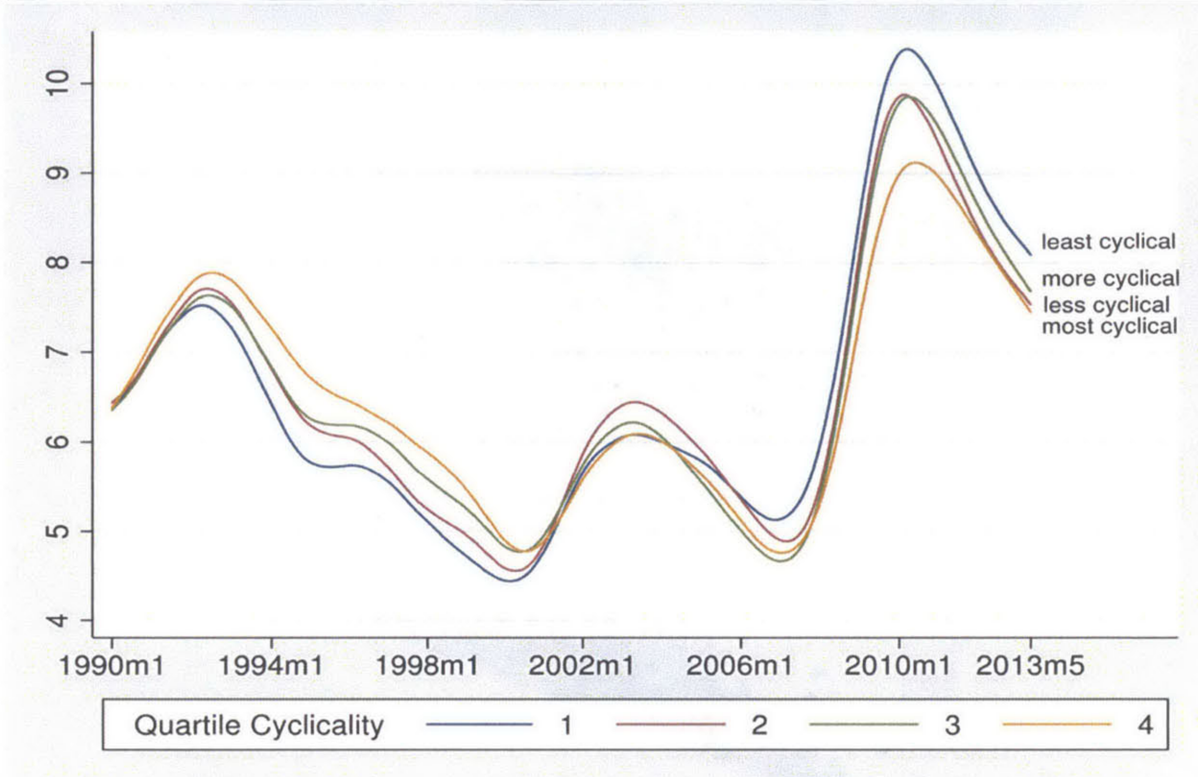
Notes: Figure plots observed log home price indices and predicted indices using long-run variation in the price cycle. The right-hand panel lines show the predicted values from a first stage regression of log(HPI) on CBSA fixed effects and the instrument set, as specified in equation (8) in the text.

Figure 7. Subprime Market Share by Long-Run Price Cyclicality Quartile



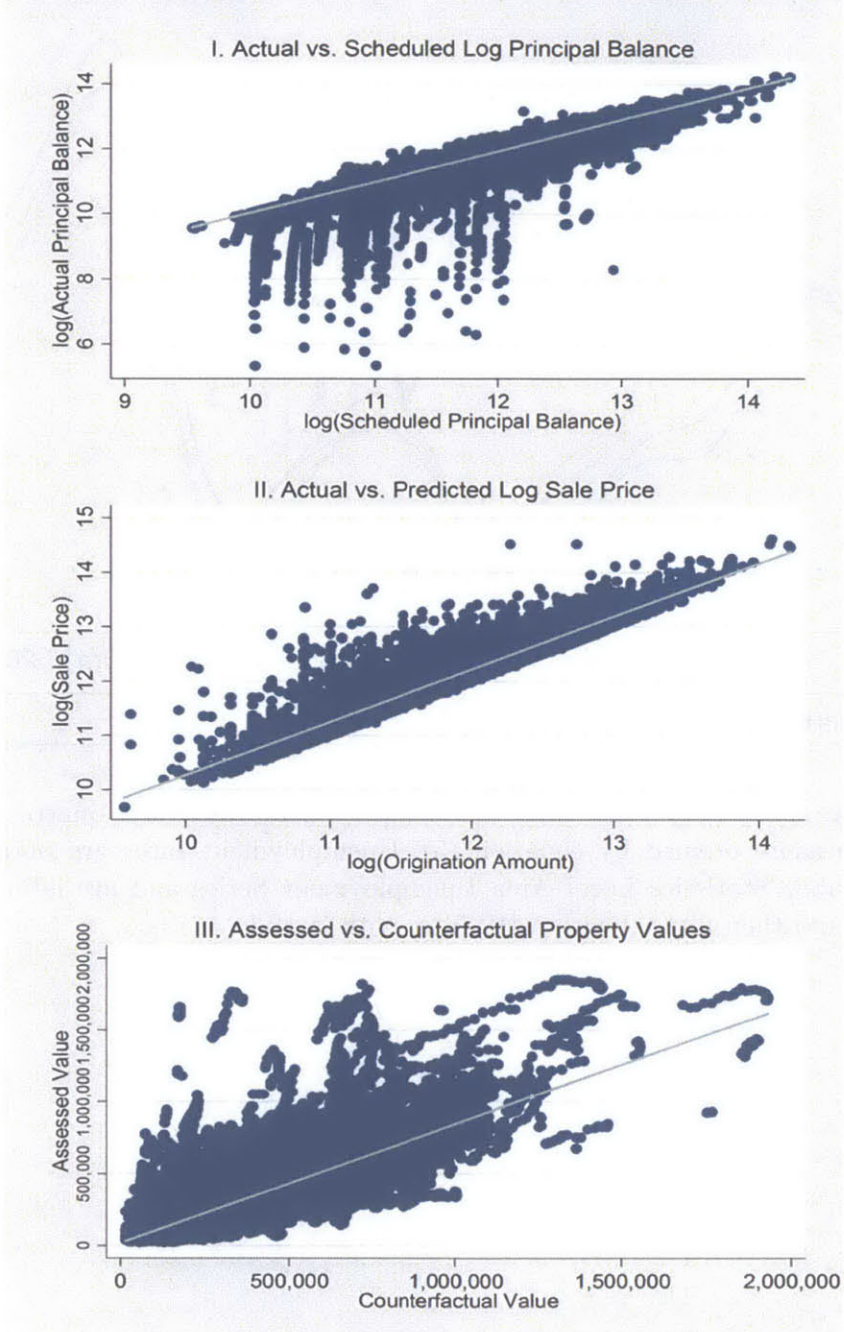
Notes: Figure shows average subprime market share by quartile of the price cyclicality measure defined by equation (5). Subprime market shares are calcualted using HMDA data as the fraction of mortgages originated in a given year that were made by a lender on the HUD subprime lender's list in any year and adjusted for CBSA fixed effects.

Figure 8. Unemployment Rates by Long-Run Price Cyclicality Quartile



Notes: Figure shows average filtered unemployment rates by quartile of the price cyclicality measure defined by equation (5). Unemployment rates are obtained from the Bureau of Labor Statistics Local Area Unemployment Series and are adjusted for CBSA fixed effects and then filtered with a HP filter with lambda = 1,600.

Figure 9. First-Stage Plots for Combined Loan-to-Value Ratio



Notes: Panel I plots actual log principal balances versus log balances corresponding to the 30-year fixed-rate mortgage amoritization schedule. Panel II plots log sale prices against log origination amounts. Panel III plots property values against counterfactual values, imputed using home price indices predicted using long-run local variation in home-price cyclicality. Diagonal lines show the fitted bivariate linear regression line.

Table 1. Summary Statistics

	mean	sd	min	max
Default	0.24	0.42	0	1
Prepaid	0.50	0.50	0	1
Censored	0.26	0.44	0	1
2004 Cohort	0.23	0.42	0	1
2005 Cohort	0.29	0.46	0	1
2006 Cohort	0.26	0.44	0	1
2007 Cohort	0.08	0.28	0	1
FICO Score	617.26	59.12	432	881
Debt-to-Income (non-missing)	0.40	0.10	0	0.9
DTI missing	0.26	0.44	0	1
Combined LTV	0.85	0.14	0	1.57
Interest Rate	7.27	1.33	1	13
Full Documentation	0.68	0.46	0	1
Owner Occupied	0.92	0.28	0	1
Cash-out Refi	0.54	0.50	0	1
Adjustable Rate	0.54	0.50	0	1
Interest-only	0.13	0.34	0	1
Balloon	0.09	0.28	0	1
Has 2nd Lien	0.20	0.40	0	1
		N = 32,172		

Notes: Default, prepaid, and censored are indicator variables for a mortgage's termination type. The remaining characeteristics are measured at time of orignation. Full documentation, owner occupied, cash-out refinance, adjustable rate, interest-only, balloon mortgage, and has second lien are all indicator variables for the given characteristic. See Section 3 in the text for more details.

Table 2. Summary Statistics by Cohort

	Cohort				
	2003	2004	2005	2006	2007
Default	0.11	0.13	0.25	0.36	0.32
	(0.31)	(0.34)	(0.43)	(0.48)	(0.47)
Prepaid	0.76	0.71	0.52	0.28	0.18
	(0.43)	(0.45)	(0.50)	(0.45)	(0.38)
FICO Score	617.00	618.15	618.59	616.08	614.33
	(61.85)	(61.15)	(59.68)	(56.48)	(54.72)
Debt-to-Income (non-missing)	0.39	0.39	0.40	0.41	0.41
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
DTI missing	0.26	0.23	0.32	0.21	0.24
	(0.44)	(0.42)	(0.47)	(0.41)	(0.43)
Combined LTV	0.83	0.84	0.86	0.86	0.84
	(0.13)	(0.13)	(0.14)	(0.14)	(0.15)
Interest Rate	7.23	6.78	6.89	7.89	8.07
	(1.29)	(1.21)	(1.18)	(1.25)	(1.39)
Full Documentation	0.71	0.70	0.68	0.67	0.68
	(0.45)	(0.46)	(0.47)	(0.47)	(0.46)
Owner Occupied	0.91	0.91	0.92	0.92	0.91
	(0.28)	(0.28)	(0.28)	(0.27)	(0.29)
Cash-out Refi	0.57	0.57	0.53	0.51	0.58
	(0.50)	(0.49)	(0.50)	(0.50)	(0.49)
Adjustable Rate	0.61	0.63	0.57	0.45	0.34
	(0.49)	(0.48)	(0.50)	(0.50)	(0.47)
Interest-only	0.03	0.11	0.21	0.13	0.09
	(0.16)	(0.31)	(0.41)	(0.33)	(0.29)
Balloon	0.01	0.00	0.02	0.21	0.28
	(0.10)	(0.04)	(0.15)	(0.41)	(0.45)
Has 2nd Lien	0.07	0.15	0.24	0.28	0.16
	(0.25)	(0.36)	(0.42)	(0.45)	(0.37)
Observations	4407	7251	9444	8336	2734

Notes: Table reports means and standard deviations in parentheses of individual loan characteristics by borrower cohort. See notes to Table 1 for further details.

Table 3. Effects of Loan Characteristics and Prices:
Default Hazard Model Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
2004 Cohort	0.217***	0.223***	0.290***	0.234***	0.188***	0.137**	0.094
	(0.072)	(0.071)	(0.073)	(0.069)	(0.069)	(0.068)	(0.066)
2005 Cohort	0.717***	0.709***	0.747***	0.564***	0.519***	0.407***	0.190***
	(0.100)	(0.097)	(0.097)	(0.089)	(0.087)	(0.075)	(0.068)
2006 Cohort	0.954***	0.984***	0.820***	0.556***	0.579***	0.470***	0.045
	(0.130)	(0.129)	(0.128)	(0.118)	(0.121)	(0.093)	(0.086)
2007 Cohort	0.734***	0.800***	0.613***	0.424***	0.466***	0.235***	-0.107
	(0.120)	(0.116)	(0.120)	(0.112)	(0.116)	(0.083)	(0.084)
12-month Δlog(HPI)						-3.685***	-3.857***
						(0.131)	(0.152)
CBSA FE	n	y	y	y	y	y	y
Borrower Characteris	n	n	y	n	y	n	y
Loan Characteristics	n	n	n	y	y	n	y
Observations	1,224,716	1,224,716	1,224,716	1,224,716	1,224,716	1,224,716	1,224,716
Log likelihood	-44,335	-43,574	-42,642	-43,186	-42,498	-43,142	-42,033

Notes: Table reports maximum-likelihood estimates of the default hazard model given in equations (2) and (3) in the text. All specifications include indicator variables for each value of loan age as a non-parametric baseline hazard. Standard errors in parentheses are clustered at the CBSA level.

Table 4. Default Hazard Model Estimates: Allowing for Unobserved Heterogeneity

	(1)	(2)	(3)	(4)	(5)
2004 Cohort	0.254***	0.188**	0.130**	0.078	0.113*
	(0.073)	(0.082)	(0.056)	(0.063)	(0.065)
2005 Cohort	0.924***	0.674***	0.372***	0.172***	0.250***
	(0.077)	(0.078)	(0.052)	(0.059)	(0.061)
2006 Cohort	1.361***	0.865***	0.380***	0.006	0.117*
	(0.096)	(0.082)	(0.052)	(0.061)	(0.064)
2007 Cohort	1.079***	0.682***	0.110*	-0.186***	-0.039
	(0.097)	(0.093)	(0.060)	(0.070)	(0.074)
12-month $\Delta\log(\text{HPI})$			-4.063***	-4.743***	-4.139***
			(0.103)	(0.128)	(0.140)
Borrower Characteristics	n	y	n	y	y
Loan Characteristics	n	y	n	y	y
State Fixed Effects	n	n	n	n	y
Observations	1,224,716	1,224,716	1,224,716	1,224,716	1,224,716
Log likelihood	-44,295	-43,140	-43,631	-42,476	-42,284

Notes: Table reports maximum-likelihood estimates of the default hazard model given in equations (2) and (7) in the text. All specifications include indicator variables for each value of loan age as a non-parametric baseline hazard. Standard errors in parentheses are homoskedastic MLE standard errors.

Table 5. Effect of Long-Run Cyclicality on Price Changes: First-Stage Results

	(1)	(2)	(3)
2004 Cohort		-0.011***	-0.012***
		(0.003)	(0.003)
2005 Cohort		-0.031***	-0.031***
		(0.007)	(0.007)
2006 Cohort		-0.053***	-0.051***
		(0.011)	(0.011)
2007 Cohort		-0.064***	-0.059***
		(0.013)	(0.012)
Baseline hazard	n	y	y
CBSA FE	n	y	y
Borrower covariates	n	n	y
Loan covariates	n	n	y
Observations	1,224,716	1,224,716	1,224,716
R-squared	0.497	0.559	0.562
Partial F-stat	49.04	31.23	30.97

Notes: Table estimates first stage specifications detailed by equation (11) by OLS. Dependent variable is the 12-month change in the log house price index. The instruments are calendar month indicator variables interacted with the historical cyclicality measure defined by equation (9) in the text. Standard errors are clustered by CBSA.

Table 6. Effect of Accounting for Endogeneity of Prices:
Control-Function Estimates of Default Hazard

	(1)	(2)	(3)	(4)	(5)	(6)
2004 Cohort	0.137**	0.127*	0.123*	0.094	0.083	0.078
	(0.068)	(0.068)	(0.068)	(0.066)	(0.066)	(0.066)
		[0.172]	[0.158]		[0.105]	[0.106]
2005 Cohort	0.407***	0.362***	0.357***	0.190***	0.142**	0.134**
	(0.075)	(0.076)	(0.075)	(0.068)	(0.068)	(0.066)
		[0.195]	[0.178]		[0.104]	[0.098]
2006 Cohort	0.470***	0.393***	0.403***	0.045	-0.034	-0.028
	(0.093)	(0.095)	(0.093)	(0.086)	(0.089)	(0.086)
		[0.242]	[0.212]		[0.079]	[0.074]
2007 Cohort	0.235***	0.147*	0.177**	-0.107	-0.195**	-0.170*
	(0.083)	(0.088)	(0.087)	(0.084)	(0.089)	(0.087)
		[0.256]	[0.203]		[0.084]	[0.076]
12-month Δlog(HPI)	-3.685***	-4.356***	-4.479***	-3.857***	-4.576***	-4.722***
	(0.131)	(0.352)	(0.325)	(0.152)	(0.362)	(0.335)
		[0.872]	[0.817]		[0.413]	[0.387]
Δlog(HPI) Fitted Residuals		0.941**	1.138***		1.004**	1.236***
		(0.431)	(0.413)		(0.448)	(0.433)
		[0.918]	[0.851]		[0.463]	[0.431]
Unemployment Rate			-0.020**			-0.021**
			(0.008)			(0.009)
			[0.022]			[0.011]
CBSA FE	y	y	y	y	y	y
Borrower Characteristics	n	n	n	y	y	y
Loan Characteristics	n	n	n	y	y	y
Observations	1,224,716	1,224,716	1,223,448	1,224,716	1,224,716	1,223,448
Log likelihood	-42,305	-43,138	-43,103	-42,033	-42,029	-41,993

Notes: Table reports maximum-likelihood control-function estimates of the default hazard model given in equations (2) and (7) in the text. Fitted residuals are estimated from a linear first stage regression of the 12-month change in the log price index on the instruments and remaining controls. All specifications include indicator variables for each value of loan age as a non-parametric baseline hazard. Standard errors in parentheses are clustered at the CBSA level. Standard errors in brackets are from 200 block bootstrap replications.

Table 7. Effect of Allowing a Flexible Endogeneity Distribution:
Nonparametric Control-Function Estimates of Default Hazard

	(1)	(2)	(3)	(4)	(5)	(6)
2004 Cohort	0.127*	0.120*	0.120*	0.083	0.077	0.077
	(0.068)	(0.067)	(0.067)	(0.066)	(0.065)	(0.065)
	[0.172]	[0.166]	[0.166]	[0.105]	[0.105]	[0.105]
2005 Cohort	0.362***	0.337***	0.339***	0.142**	0.117*	0.120*
	(0.076)	(0.070)	(0.071)	(0.068)	(0.064)	(0.064)
	[0.195]	[0.196]	[0.198]	[0.104]	[0.108]	[0.109]
2006 Cohort	0.393***	0.358***	0.359***	-0.034	-0.066	-0.065
	(0.095)	(0.089)	(0.089)	(0.089)	(0.086)	(0.087)
	[0.242]	[0.242]	[0.244]	[0.079]	[0.081]	[0.082]
2007 Cohort	0.147*	0.122	0.121	-0.195**	-0.219**	-0.219**
	(0.088)	(0.087)	(0.087)	(0.089)	(0.091)	(0.092)
	[0.256]	[0.257]	[0.259]	[0.084]	[0.084]	[0.084]
12-month Δlog(HPI)	-4.356***	-4.737***	-4.658***	-4.576***	-4.944***	-4.877***
	(0.352)	(0.402)	(0.403)	(0.362)	(0.415)	(0.415)
	[0.872]	[0.946]	[0.970]	[0.413]	[0.459]	[0.465]
Δlog(HPI) Fitted Residuals	0.941**	0.263	0.615	1.004**	0.413	0.699
	(0.431)	(0.494)	(0.552)	(0.448)	(0.494)	(0.567)
	[0.918]	[1.045]	[1.093]	[0.463]	[0.543]	[0.558]
(Δlog(HPI) Fitted Residuals)^2		-4.513***	-1.215		-4.841***	-1.764
		(1.580)	(4.205)		(1.791)	(4.238)
		[2.710]	[6.134]		[1.854]	[4.071]
(Δlog(HPI) Fitted Residuals)^3		23.690**	-7.211		19.707**	-4.513
		(9.852)	(32.970)		(8.849)	(28.217)
		[10.317]	[28.550]		[8.665]	[24.850]
(Δlog(HPI) Fitted Residuals)^4			-75.177			-71.583
			(65.940)			(64.402)
			[93.464]			[77.720]
(Δlog(HPI) Fitted Residuals)^5			282.743			195.488
			(446.678)			(381.055)
			[394.651]			[356.268]
CBSA FE	y	y	y	y	y	y
Borrower Characteristics	n	n	n	y	y	y
Loan Characteristics	n	n	n	y	y	y
Observations	1,224,716	1,224,716	1,224,716	1,224,716	1,223,448	1,224,716
Log likelihood	-43,138	-43,114	-43,111	-42,029	-42,007	-42,006

Notes: See Table 5 notes.

Table 8. Effect of Current Combined Loan-to-Value Ratio on Default Hazard: Control Function Results

	(1)	(2)	(3)	(4)	(5)	(6)
2004 Cohort	0.202***	0.188***	0.127**	0.116*	0.110*	0.092
	(0.066)	(0.064)	(0.063)	(0.064)	(0.064)	(0.063)
2005 Cohort	0.424***	0.348***	0.138**	0.126*	0.133**	0.095
	(0.075)	(0.068)	(0.063)	(0.066)	(0.064)	(0.063)
2006 Cohort	0.317***	0.159*	-0.185**	-0.162**	-0.125	-0.146*
	(0.096)	(0.083)	(0.076)	(0.078)	(0.079)	(0.077)
2007 Cohort	0.143	-0.055	-0.417***	-0.372***	-0.292***	-0.279***
	(0.096)	(0.084)	(0.078)	(0.081)	(0.089)	(0.085)
Underwater	0.683***			0.284***		
	(0.060)			(0.052)		
12-month $\Delta\log(\text{HPI})$			-3.221***	-4.693***	-4.718***	-4.814***
			(0.237)	(0.333)	(0.382)	(0.371)
CLTV Fitted Residuals				0.007***	0.013***	0.011***
				(0.000)	-0.001	(0.001)
$\Delta\log(\text{HPI})$ Fitted Residuals				1.579***	2.252***	2.170***
				(0.406)	(0.500)	(0.495)
Unemployment Rate						-0.050***
						(0.013)
CLTV Linear Spline	n	y	y	n	y	y
Observations	1,037,581	1,037,581	1,037,581	1,036,611	1,037,581	1,036,611
Log likelihood	-35,935	-35,723	-35,477	-35,444	-35,379	-35,329

Notes: Table reports maximum-likelihood estimates of the default hazard model given in equations (2) and (14) in the text. Current combined loan-to-value ratios (CLTVs) are calculated by LoanPerformance as the total outstanding principal on a loan divided by an automated assessing model's estimate of the market value of each home. Underwater is an indicator for CLTV>1. The linear spline is defined by equation (15) in the text. All specifications include individual loan and borrower characteristics, CBSA fixed effects, and indicator variables for each value of loan age as a non-parametric baseline hazard function. Standard errors in parentheses are clustered at the CBSA level.

Table 9. First-Stage Results for Combined Loan-to-Value Ratios

	(1)	(2)	(3)	(4)	(5)
log(Origination Amount)	-0.640***	-0.724***	-0.950***	-0.958***	-0.968***
	(0.047)	(0.043)	(0.047)	(0.046)	(0.046)
log(Prinicipal Balance)	0.787***	0.917***	1.063***	1.066***	1.076***
	(0.044)	(0.042)	(0.047)	(0.046)	(0.046)
Predicted Appreciation	-1.265***	-0.627***	-0.610***	-0.602***	-0.103
	(0.162)	(0.137)	(0.138)	(0.137)	(0.100)
2004 Cohort		-0.001	0.003	0.001	0.019***
		(0.008)	(0.007)	(0.007)	(0.007)
2005 Cohort		0.030**	0.038***	0.024**	0.072***
		(0.013)	(0.013)	(0.012)	(0.013)
2006 Cohort		0.091***	0.110***	0.084***	0.146***
		(0.021)	(0.022)	(0.020)	(0.023)
2007 Cohort		0.156***	0.173***	0.145***	0.190***
		(0.028)	(0.028)	(0.026)	(0.027)
12-month Δlog(HPI)				-0.326***	-0.204***
				(0.024)	(0.032)
Unemployment Rate					0.052***
					(0.006)
Baseline hazard	n	y	y	y	y
CBSA FE	n	y	y	y	y
Borrower covariates	n	n	y	y	y
Loan covariates	n	n	y	y	y
Observations	1,037,581	1,037,581	1,037,581	1,037,581	1,036,611
R-squared	0.242	0.355	0.423	0.428	0.462
Partial F-stat	239.10	331.70	232.29	231.66	221.20

Notes: Table estimates first stage specifications detailed by equation (16) by OLS. Dependent variable is current combined loan-to-value ratio, calculated by CoreLogic as the total outstanding principal on a loan divided by an automated assessing model's estimate of the market value of each home. Standard errors are clustered by CBSA.

Table 10. Counterfactual Annual Default Rates by Cohort

	Counterfactual Scenario		Default Rate by Cohort					
	Prices	Loan Characteristics	2003	2004	2005	2006	2007	Overall
(1)	Actual	Actual	4.2%	5.3%	9.2%	12.0%	9.8%	8.7%
(2)	Actual	2003	3.7%	4.6%	7.0%	7.4%	6.5%	6.2%
(3)	Actual	2006	5.3%	6.6%	9.9%	10.5%	9.2%	8.8%
(4)	2003	Actual	4.1%	4.6%	5.6%	5.6%	4.4%	5.1%
(5)	2006	Actual	8.5%	9.4%	11.3%	11.0%	8.4%	10.2%
(6)	No price change	Actual	6.3%	6.9%	8.2%	7.9%	6.0%	7.4%
(7)	No price change	2003	5.3%	5.7%	6.0%	4.7%	3.9%	5.3%
(8)	No price change	2006	7.6%	8.1%	8.5%	6.8%	5.6%	7.5%
		Observations	115,567	193,554	281,346	285,277	106,764	982,508

Notes: Table reports estimated annual default rates under the indicated counterfactual scenarios for prices and loan characteristics. Annual default rates are defined as 12 times the average fraction of loans that default in each month, measured over all existing loans within five years of origination. Scenarios using actual characteristics retain observed covariates. Scenarios using a given year's prices replace all price changes with the average price changes faced by the given year's borrowers at each value of loan age. Scenarios using a given year's loan characteristics assign all loans the average characteristics from the indicated cohort.

Chapter 2

Housing Market Spillovers: Evidence from the End of Rent Control in Cambridge Massachusetts

with David H. Autor and Parag A. Pathak

Introduction

Spillovers from the attributes and actions of neighborhood residents onto the value of surrounding properties and neighborhoods are central to the theory of urban economics and the development of efficient housing policy (Fujita, 1991; Glaeser and Gyourko, 2009). Credibly identifying and quantifying these external effects, however, poses a significant empirical challenge because key features of the housing market equilibrium—in particular, who lives where, the quality and quantity of housing, the levels of local public goods and amenities, and what prices prevail—are all determined simultaneously in equilibrium.[1]

This paper exploits an unusual, large scale policy change, the elimination of rent control in Cambridge, Massachusetts in 1995, to quantify the capitalization of residential housing market externalities onto the value of residential real estate. From December 1970 through 1994, all rental units in

[1]See Kasy (2013) for a recent discussion of the nonparametric identification in location choice models with social externalities.

Cambridge built prior to 1969 were regulated by a far-reaching rent control ordinance that placed strict caps on rent increases and tightly restricted the removal of units from the rental stock. The legislative intent of the rent control ordinance was to provide affordable rental housing. and at the eve of rent control's elimination in 1994, controlled units typically rented at 40-plus percent below the price of nearby non-controlled properties—though maintenance and amenities in controlled units tended to be sub-par (Sims, 2007).[2]

The policy change that provides the identifying variation for our study is the swift elimination of Cambridge's rent control law via a state-wide ballot initiative. In November 1994, the Massachusetts electorate passed a referendum to eliminate rent control by a narrow 51% to 49% margin, with nearly 60% of Cambridge residents voting to retain the rent control ordinance. Thus, rent decontrol in Cambridge, which commenced only two months after the November 1994 referendum, was voted into law by Massachusetts cities and towns that had never experienced rent control, while ironically, the three Massachusetts municipalities with active rent control regimes—Cambridge, Boston, and Brookline, each of which voted to maintain rent control—were overruled by the state-wide majority.[3]

Alongside its swift and largely unanticipated elimination, two unusual features of Cambridge's rent control ordinance make it well suited to credibly identify the effects of rent control on residential housing markets. First, because the rent control ordinance only applied to a fixed, non-expanding set of residential units—specifically, non-owner occupied rental houses, condominiums, or apartments built prior to 1969—controlled and never-controlled units stood side by side in Cambridge neighborhoods on the eve of rent control removal, thus offering a tight temporal and geographic framework for assessing the impact of the law on residential property prices.[4] Second, although roughly a third of residential units were controlled prior to elimination (see Figure 1), this fraction frequently exceeded sixty percent in neighborhoods that had older housing stocks and a substantial share of renters at the time of rent control's enactment in 1970. This sizable cross-neighborhood variation allows us to assess localized price effects by comparing pre- and post-removal price appreciation among both decontrolled and never-controlled properties in neighborhoods that differed in their 'rent control intensity'—that is, the share of residential units that were controlled.

[2]Using microdata from a 1987 Abt Associates study commissioned by the City of Cambridge (Finkel and Wallace, 1987), we estimate that quality-adjusted rents were approximately 44 percent lower at controlled units than at observably similar non-controlled units.

[3]As discussed in Sims (2007), the Boston and Brookline rent control regimes were far less comprehensive than in Cambridge.

[4]If an owner-occupied residential unit built before 1969 were put up for rent, it could be subject to rent control. Our informal understanding based on discussions with Cambridge homeowners of that era was that such rentals were rare and often arranged discreetly to avoid the notice of the Rent Control Board.

Our conceptual model and empirical work distinguish two channels through which rent decontrol may affect the market values of residential properties. The first, which we term the direct effect, reflects the capitalization of landlords' newfound ability to charge market rents. Absent any change in residential investments or neighborhood characteristics—and assuming that price controls were binding—rent control removal should directly raise the ownership value of formerly controlled properties by uncapping rents and, simultaneously, increasing the returns to landlord investments. The second channel, which we term the indirect effect, encompasses the multiple complementary mechanisms by which rent decontrol may affect the desirability of surrounding properties: owners renovate and modernize decontrolled units, raising their rental values; affluent tenants who particularly value these amenities rent these units as incumbents depart in the face of rising prices; higher income tenants move into nearby never-controlled properties, attracted by the amenities of an improved housing stock and more affluent neighbors; property owners make further investments in both decontrolled and never-controlled units as overall tenant income levels rise.[5]

Distinct from the 'direct' effect of decontrol, which by definition operates only on formerly controlled properties, the indirect channel may affect the market value of both decontrolled and never-controlled properties by increasing the desirability of the neighborhoods in which they're located. While our analysis does not allow us to further decompose the indirect effect into its constituent components (investment, reallocation, and the complementarities between the two), historical evidence suggests that each of these channels was relevant. Because Cambridge's Rent Control Board was unlikely to grant rent increases following property improvements, it was widely perceived that rent control muted owners' incentives to maintain and improve controlled properties.[6] Consistent with this view, Sims (2007) finds that chronic maintenance problems—such as holes in walls or floors, chipped or peeling paint and loose railings—were more prevalent in controlled than non-controlled units during the rent control era, and that this differential fell substantially with rent control's elimination. The end of rent control also spurred substantial tenant turnover. Cambridge's rent control law was intended to enable less affluent tenants to reside in units that would command high rents under a market allocation, particularly the dense neighborhoods proximate

[5] Increases in residential investment after rent decontrol do not divide cleanly into direct or indirect effects: absent spillovers, decontrol should raise the return to renovations and repairs of ill-maintained decontrolled units (a direct effect); the complementarities among tenant incomes, neighborhood amenities, and the quality of the housing capital stock should raise the return to investments at both decontrolled and never-controlled units (an indirect effect).

[6] Leonard (1981) notes that the Board limited the allowable rate of return on investments at a "relatively low" level deemed "fair," which made improvements both comparatively unprofitable and difficult to finance. Rent Control Board records indicate that applications for rent adjustments were infrequent—once per decade for a typical unit.

to Cambridge's major universities, commercial centers, and transportation hubs. While there was no formal mechanism to allocate controlled units to low-income households, limited quantitative evidence indicates that less affluent residents and students were overrepresented in controlled units—though a significant number of units were also occupied by wealthy professionals.[7] As we show below, exit rates from formerly controlled units spiked in the years immediately following rent decontrol. And given the substantial accompanying increases in rents, it is likely that the new cohorts of renters were significantly more affluent than the tenants they replaced. Our analysis will capture the net effect of these potentially mutually reinforcing channels on the market value of Cambridge residential real estate.

Regulations are widespread in housing markets, and rent controls are arguably among the most important historically (Friedman and Stigler, 1946; Glaeser and Gyourko, 2009). Because they directly manipulate the price mechanism, they are likely to reshape the allocation of residents to locations, the incentives for investment and maintenance of controlled units, and the supply, demand, quality and allocation of units in the non-controlled sector. The modern era of U.S. rent controls began as a part of World War II-era price controls and as a reaction to housing shortages following demographic changes immediately after the war (Fetter, 2013). While the prevalence of rent control as a housing market policy has decreased since this period, rent control and rent stabilization plans are still in place in many U.S. and European cities (Arnott, 1995). New York City's system of rent regulation affects at least one million apartments, while cities such as San Francisco, Los Angeles, Washington DC, and several California and New Jersey cities have various forms of rent regulation. Rent control remains a topic of active debate among affordable housing advocates.

The early empirical literature on rent control focuses on its effects on the supply of housing services (Olsen, 1972) and the incentives of landlords to invest in building quality (Frankena, 1975; Gyourko and Linneman, 1989). A second strand of this literature examines how below-market rents may encourage individuals to spend effort to obtain cheap housing, leading to a misallocation of housing (Suen, 1989; Glaeser and Luttmer, 2003; Sims, 2011). Fallis and Smith (1984) examine how the impact of rent control on the uncontrolled sector depends on the allocation mechanism in the

[7] A 1998 study commissioned by the City of Cambridge found that sitting residents of formerly controlled units had mean annual earnings in 1997 of $35,650 versus $43,630 among tenants of market rate units and $41,340 among tenants of formerly controlled units who had taken residence after rent control removal (Atlantic Marketing Research, 1998). Sims (2007) calculates that 67 percent of residents of rent controlled units in Boston, Brookline, and Cambridge were in the bottom two quartiles of the income distribution. At the same time, blacks were substantially underrepresented in controlled units.

controlled sector. Wang (2011) investigates the impact of privatization of housing that was owned and allocated by the state in urban China. Her analysis, like ours, shows that the degree of misallocation of assets prior to privatization impacts the expected change in prices.

Sims (2007) undertakes the first empirical analysis of the end of rent control in Massachusetts, exploring its impacts on the supply of rental properties and their rental prices. Sims shows that the elimination of rent control spurred substantial rent increases in Massachusetts towns that had binding rent control laws in 1994 (Boston, Brookline and Cambridge) and led to significant increases in the quality and quantity of rental housing available. Distinct from Sims' work, we analyze rent control's effect on the market value (rather than rental prices) of the entire residential housing stock (not simply rental units) in Cambridge and distinguish its effects on decontrolled properties and never-controlled properties.[8]

Our work is also related to studies of neighborhood revitalization and gentrification, both of which may generate spillover benefits to surrounding areas (Hurst, Guerrieri and Hartley, 2013; Ioannides, 2003; Rossi-Hansberg, Sarte and Owens, 2010; Schwartz, Ellen, Voicu and Schill, 2006). Studies by Linden and Rockoff (2008) and Pope (2008) of the housing market impacts of the arrival of registered sex offenders into a neighborhood consider allocative externalities in residential housing. Recent interest in measuring external effects in housing has been spurred in part by historically high levels of foreclosures and the concern for their impact on immediate neighbors and neighborhoods (Campbell, Giglio and Pathak, 2011; Hartley, 2010; Mian, Sufi and Trebbi, 2011).[9]

Our analysis draws on a uniquely detailed geographic and economic database sourced from Cambridge administrative record that enumerates the exact location of all rent controlled units, the assessed value of each house and condominium in 1994 and 2004, the transacted price of each residential property sold between 1988 and 2005, the movement of properties across various residential and non-residential uses (e.g., houses that were converted to condominiums), and the permitted investment expenditures at each residential location. We additionally use 10 years of Cambridge City Census data to document the rapid turnover of residents of formerly controlled units following the end of rent control. These sources permit direct estimation of changes in residential real estate prices induced by rent decontrol.

[8] Sims (2007) further explores spillovers from decontrol onto the rental price of never-controlled units, but his data do not allow sufficient precision to draw firm conclusions.

[9] In addition, a number of papers present evidence that subprime mortgage lending may lead to price appreciation in neighborhoods where housing credit was historically in short supply (Mian and Sufi, 2009; Landvoigt, Piazzesi and Schneider, 2012).

We find compelling evidence that the elimination of rent control raised the market values of both decontrolled and never-controlled properties. Our main estimates imply that during the rent control era, rent controlled properties were valued at a discount of about 45 to 50 percent relative to never-controlled properties with comparable characteristics in the same neighborhoods, and that their assessed values rose by 18 to 25 percent relative to never-controlled properties following rent decontrol. This differential appreciation should primarily reflect the direct effect of rent decontrol on the market value of formerly controlled units generated by the potential for owners to charge market rents, the option to convert rental units into condominiums, and the flow of returns from associated capital investments.

To assess whether rent control density affected the desirability of neighborhoods over and above its direct effect on controlled properties, we next calculate a rent control exposure measure for each residential unit that is equal to the fraction of other residential units within a 0.20 mile radius that were subject to rent control as of 1994. A central finding is that post-decontrol price appreciation was significantly greater at units that had a larger fraction of formerly controlled neighbors: residential properties at the 75^{th} percentile of rent control exposure gained approximately 13 percent more in assessed value following decontrol than did properties at the 25^{th} percentile of exposure. This differential appreciation of properties in rent control intensive locations was equally pronounced among decontrolled and never-controlled units, suggesting that rent control removal spurred overall gains in neighborhood desirability.

These findings are robust to many alternative measures of rent control intensity, to rich controls for property-level characteristics (such as age, lot size, and number of bedrooms and bathrooms), and to the inclusion of detailed geographic fixed effects and neighborhood trends that allow price levels to vary across Cambridge neighborhoods and to trend over time within them. Data on transaction prices for all properties sold in Cambridge between 1988 - 2005, which provide an alternative source for measuring changes in market values, yield comparable estimates of spillover effects to those found using the assessor's data.

One channel through which the removal of rent controls may have raised Cambridge housing values is by spurring additional capital investments. Using administrative data on residential expenditures permitted by the Cambridge Inspectional Services, we find that aggregate annual permitted building expenditures increased dramatically for both houses and condominiums after 1994, rising from \$21 million per year between 1991-1994 to \$45 million per year between 1995 and 2004. Moreover, the incidence of permitting—though not investment expenditures per unit—rose differentially at formerly controlled properties in the years immediately following rent control removal. But the *total* value of Cambridge residential investments in these 10 years was less than one quarter as large as the estimated *increment* to Cambridge residential housing values induced by rent control removal, suggesting that the allocative rather than the investment channel is the more important explanation for the post-1994 rise in the market value of never-controlled properties.

The economic magnitude of the effect of rent control removal on the value of Cambridge's housing stock is large, contributing \$2.0 billion of \$7.7 billion in Cambridge property appreciation in the decade between 1994 and 2004. Of this total effect, only \$300 million is accounted for by the direct effect of decontrol on formerly controlled units (holding exposure constant), while \$1.7 billion is due to the indirect effect. Notably, the majority of this indirect effect (\$1.1 of \$1.7 billion) stems from the differential appreciation of never controlled units. Combining both direct and indirect effects, our estimates imply that more than half (55 percent) of the capitalized cost of rent control was borne by owners of never-controlled properties.

The paper proceeds as follows. Section 2.1 provides additional detail on the enactment, enforcement, and removal of rent control in Cambridge. Section 2.2 describes a simple model of housing markets in the presence of rent control to guide our empirical analysis (the Theory Appendix contains the model). Section 2.3 describes data sources and our empirical strategy. Section 2.4 presents our main results using property assessments, while Section 2.5 presents results on the time path of the capitalization of rent decontrol using transaction prices. Section 2.6 reports on our investigation of permitting and investment activity and Section 2.7 considers economic magnitudes. We conclude with a discussion of areas for further investigation.

2.1 Cambridge Rent Control: Enactment, Enforcement, and Removal

2.1.1 Rent control adoption and elimination

In 1970, the Massachusetts state legislature enacted a statute allowing cities and towns with populations over 50,000 to implement rent control to "alleviate the severe shortage of rental housing..."[10]

[10] Quoted from "An Act Enabling Certain Cities and Towns to Control Rents and Evictions", 1970 Mass. Acts 842.

Boston, Brookline, Cambridge, Lynn, and Somerville each adopted a rent control plan, with Cambridge moving first in 1970 and keeping the ordinance longer than any other city. Lynn repealed its plan in 1974 and Somerville in 1979. Boston allowed for decontrol of vacant units in 1976 and Brookline began to phase out its system prior to the state-wide repeal, though both cities still had a significant number of controlled units in 1994 (Cantor, 1995).[11] In Cambridge, rent control was seen as an integral part of the city's affordable housing program.

Cambridge's initial rent control policy adopted in 1970 applied to all non-owner-occupied rental housing built before 1969. It did not apply to structures built after January 1, 1969, to owner-occupied condominiums, or to non-residential structures converted to rental properties after this time. Oversight of the rent control law rested with the Cambridge Rent Control Board, whose official charter was to ensure that landlords obtained a fair net operating income. The Board established maximum allowable rents for each controlled property with the aim of fixing landlord net operating income at inflation-adjusted 1967 levels. In the 1970s and 1980s, the Board authorized a series of across-the-board rent increases ranging from 1.15 to 3.1 percent, intended to cover increases in heating costs, operating costs, and property taxes (*Rent Control Board*, 1982). Landlords could also apply to raise prices above the scheduled increases, but these variances were rarely sought or granted in practice, in part because the application required supporting petitions, extensive legal documentation, and significant time investment.[12]

Distinct from many cities, Cambridge's rent control policy did not allow for so-called "vacancy decontrol," whereby controlled rental units were returned to market-rate rents after protected tenants moved out. Landlords therefore faced an incentive to remove units from the rental stock, which they accomplished by converting substantial numbers of rental units to condominiums and selling them to owner-occupants. To prevent the controlled rental stock from being depleted, the city council passed in 1979 the "Removal Permit Ordinance," which substantially restricted the removal of controlled units from the rental stock and complicated the conversion of controlled units into owner-occupied condominiums.[13]

The development that ultimately lead to rent control's elimination was the Cambridge Small Prop-

[11] See Epple (1988) for a game-theoretic model of communities' decisions to adopt rent control.

[12] A legendary incident involves Harvard Philosophy Professor Robert Nozick extracting a settlement of over $30,000 in the 1980s from his landlord, famed classicist and novelist Eric Segal, for overcharging rent, described in Tucker (1986).

[13] This ordinance required proof that removal would not aggravate the housing shortage and would "benefit the persons sought to be protected" by the rent control statute (Cantor, 1995). The ordinance was subsequently amended following difficulties with enforcement, which were made salient by the fate of the so-called "condo martyrs"—owners who were prosecuted for occupying their own controlled properties before the completion of a conversion.

erty Owners Association's successful effort to place rent control on the state-wide ballot in 1994. Putting rent control to a state-wide vote diluted the strong support that rent control enjoyed in the three municipalities with extant rent control ordinances (Boston, Brookline, and Cambridge). Rent control was eliminated by a slim 51 to 49 percent margin in November 1994, despite nearly 60 percent of Boston, Brookline, and Cambridge voters voting to retain the current regime. Just two months later in January of 1995, a majority of properties were decontrolled. A last-minute legislative compromise, however, allowed disabled, elderly, and low-income renters to retain their current units at their controlled rents for up to two years. Though only a small share of residents received rent control extensions, this compromise likely created some uncertainty about whether decontrol was final—at least until the grandfathering period expired in 1997 with no further controls in place.[14]

2.1.2 The post decontrol regime

The elimination of rent control catalyzed a series of rapid changes in the Cambridge rental market, beginning with rising rents. A 1998 survey commissioned by the City of Cambridge (Atlantic Marketing Research, 1998) found that nominal Cambridge median rents rose by 40 percent between 1994 and 1997 for tenants of formerly controlled units who either remained at these units or moved to other non-controlled units. Median rents rose by only 13 percent for sitting tenants of never-controlled units in the same time period.

Rising rents spurred a sharp increase in resident turnover at formerly controlled units after 1994, which we document by constructing a panel of all Cambridge adults ages 17-plus by street address using city voter registration records for the years 1991 through 2000.[15] On average, 26.9 percent of Cambridge residents changed locations annually, with the highest turnover rates found among apartment residents (33.5 percent), followed by residents of condominiums (29.7 percent) and houses (23.2 percent).

[14] Shortly after the referendum, the state legislature adopted a bill extending rent control for five years. The Governor vetoed this bill and later signed an alternative on January 3, 1995 that granted rent control extensions of one year (two years if the rental building had more than 12 units) to renters whose incomes were below 60% of the median for the Boston MSA (or 80% of the MSA median for disabled and elderly renters). Sims (2007) reports that about 3,000 of approximately 21,000 tenants applied for exemptions while Haveman (1998) reports that 9.4% of tenants were eligible to apply.

[15] State law (M.G. L. ch. 51.4) requires an annual listing of all adult residents for voter registration, regardless of voter status, including name, street address, gender, date of birth, occupation, and nationality. City Census books from 1991-2000 were double-entry hand keyed and assembled into a panel using name and address matching, as described in the Data Appendix.

We assess whether turnover rates at formerly controlled units rose differentially after 1994 by fitting linear probability models of the following form:

$$\text{NEW}_{ijt} = \gamma_g + \delta_t + \lambda_1 \text{RC}_j + \lambda_2 \text{RC}_j \times \text{Post}_t + \epsilon_{ijt}, \tag{2.1}$$

where NEW_{ijt} is an indicator equal to one if resident i in unit j in year t was not present in that unit in the prior year. In this model, RC_j is an indicator equal to one if unit j was rent controlled in 1994, γ_g is a vector of 1990 Census block group dummies, δ_t is a vector of year dummies, and Post_t is an indicator for years 1995 onwards. Prior to 1995, residents of controlled units were not significantly more likely to turnover than residents of non-controlled units.[16] Following decontrol, the turnover differential between formerly controlled and never-controlled units rose by 5.4 percentage points, with an even larger increase at condominiums (Table 1).

Figure 2 depicts both the evolution of this turnover differential using a variant of equation (2.1) in which the rent control indicator is interacted with a set of year dummies. Turnover rates at decontrolled units spiked by 4 percentage points relative to never-controlled units in the first year of decontrol and continued to climb to 10 percentage points over the next three years. Thus, the process of resident reallocation and neighborhood change spurred by decontrol took multiple years to unfold. Interestingly, Figure 2 also shows that turnover rates at never-controlled units changed little following decontrol.

A sharp increase in residential property investments also followed the end of rent control. The number of building permits issued per residential unit for improvements and new construction increased by approximately 20 percent after 1994, and annual permitted expenditures roughly doubled in real terms (Table A1). Elimination of the "Removal Permit Ordinance" allowed a substantial number of decontrolled houses, apartments, and non-residential units to be converted to condominiums. From 1994 to 2004, Cambridge's stock of residential houses decreased by 6 percent, while the stock of condominiums increased by 32 percent, with 45 percent of this increase accounted for by conversion of houses to condominiums (Table A2).[17] At the same time, the

[16]Subsequent columns reveal that this result is driven by composition. Focusing only on apartments and condominiums, residents of controlled units were significantly less likely to turn over than residents of non-controlled units—consistent with the idea that controlled units had scarcity value. Residents of controlled houses, by contrast, were significantly more likely to turnover than residents of non-controlled houses, but this likely reflects the fact that most non-controlled houses were owner-occupied whereas controlled houses were renter-occupied.

[17]These calculations use the Cambridge Assessor's databases from 1995 and 2005, reflecting the status of properties in 1994 and 2004, respectively. We count each unit in multi-family houses separately to meaningfully compare the supply of housing across different structure types and in different periods. The stock of units in houses in Cambridge

fraction of residential units available as rental properties rose by six percentage points (Sims, 2007)

This combination of sizable rent increases, rapid turnover of incumbent renters, rising residential investment, and outward shifts in the supply of both condominiums and rental properties were likely in net to have changed the quality of the Cambridge residential housing stock, the allocation of residents to neighborhoods, and the availability of residential units for both rent and sale.

2.2 The Direct and Indirect Effects of Rent Control

The Theory Appendix presents a stylized model of the housing market, summarized here, that considers the relationship between rent control and prices of both controlled and non-controlled properties. In the model, a city consists of N neighborhoods with a continuum of locations in each neighborhood. Potential residents choose locations to maximize utility defined over consumption of housing services, a non-housing composite good, and local amenities. Residents have identical preferences and differ only in their income levels. Profit-maximizing landlords choose the level of maintenance at each location, and this level is increasing in the price of housing services.

We assume that amenities in a neighborhood depend on the housing maintenance levels and the income distribution of residents in the neighborhood, where higher maintenance and higher income neighbors are also more desirable and hence contribute more to neighborhood amenities. This formulation creates positive feedback from the extent of maintenance, residents' income, and neighborhood amenities. In the free market equilibrium (absent rent controls), rents are higher in neighborhoods with greater amenities due to higher maintenance and the presence of higher income neighbors.

We consider the imposition of rent controls at the initial free market equilibrium by assuming that a rent control authority caps the rent of some units in a neighborhood at below their free market level. Since landlords choose maintenance levels facing a regulated price, maintenance levels and hence housing services are lower at controlled units. The combination of reduced rents and lower maintenance has one of two effects on incumbent residents: either they are sufficiently compensated by reduced rents so that they remain at their current locations, although the bundle of maintenance and amenities is not optimized for their income levels; or alternatively, they choose to relocate to areas with higher amenities and higher rents. In the latter case, they will be replaced by residents

decreased from 14,722 in 1994 to 13,861 in 2004 and the stock of condominiums rose from 7,220 to 9,561 units.

who prefer lower housing services, i.e., those with lower incomes.[18] The average income at controlled locations therefore weakly declines following the imposition of rent control.

Since neighborhood amenities are a function of the maintenance of *all* units in a neighborhood and the neighborhood income distribution, the level of amenities at *non-controlled* locations in these neighborhoods—as well as maintenance and rents—are also impaired by rent control. This in turn causes lower income residents to move into non-controlled locations. Thus, rent control causes inefficiently low maintenance and misallocation of residents at both controlled *and* non-controlled locations within a neighborhood.

Decontrol unwinds these effects. Prices rise due directly to the lifting of the cap, and indirectly due to improved maintenance and increased production of local amenities throughout the neighborhood. At *non-controlled* locations, the price increase will be greater in neighborhoods where a larger fraction of locations were controlled, where the capped price ceiling was set further below the market price level, and where controls induced larger resident misallocation relative to the free market setting. The lifting of controls allows an additional, direct price increase at formerly controlled locations.

The model also offers a simple welfare interpretation of any direct and indirect price effects of rent decontrol. Price increases at *decontrolled* locations reflect three forces: a mechanical 'uncapping' effect, which reflects a transfer from renters to owners; a price increase reflecting improved maintenance, which generates increased landlord surplus net of the resource cost of maintenance; and a price increase reflecting greater neighborhood amenities due to improvements in maintenance and changes in resident types nearby. While the latter two effects reflect economic gains, the first does not. The price increase at decontrolled locations is therefore likely to substantially exceed the economic gains from decontrol at these locations.

Induced price increases at *non-controlled* locations following decontrol reflect the capitalization into house values of two of these three forces: improved maintenance (or, more generally, housing investments), and greater neighborhood amenities (both due to sorting and capital improvements at other properties). Therefore, the increase in prices at non-controlled locations, net of the additional resource costs expended on maintenance and improvements, can be used to assess the external effects of decontrol—that is, the spillovers. We quantify these spillovers below by estimating the increase

[18]If the incumbent renter is dissatisfied with the new price-services pair, this pair can only be preferred by a lower type.

in market value of never-controlled units and netting out the components plausibly attributable to investment.

2.3 Data and Measurement

We briefly discuss our data sources and measurement of rent control intensity in this section, with further details in the Data Appendix.

2.3.1 Cambridge real estate

There are approximately 15,000 taxable parcels of land in the city of Cambridge organized into unique geographic units known as "map-lots." The foundation for our dataset is a snapshot of the entire universe of residential real estate from the 1995 Cambridge Assessor's File, from which we construct the residential housing structures file.[19] Each record includes the map-lot identifier, address, owner's name and address, usage, and property tax assessment as of January 1994. Usage categories are designated as commercial or residential, and residential categories are further subdivided into condominiums, single-family, two-family, and three-family houses, multi-unit apartment complexes, and mixed residential-commercial structures. In calculating rent control intensity below, we treat any usage code where individuals are likely to live as a residential structure. Our analysis of assessed values and transactions is limited to houses and condominiums, which comprise the market for residential real estate.

We identify rent controlled properties from historical records of the Cambridge Rent Control Board obtained via a Freedom of Information Act (FOIA) request.[20] We merge rent control structures to the Assessor's file using the map-lot identifier and address information coded in the Rent Control Board file. Rent controlled records that could not be matched via map-lot identifiers were hand-matched to the corresponding street address. Due to limitations of the Rent Control Board data, it was often not possible to determine which specific units in a multi-unit building were controlled. This creates a potential econometric pitfall: if we were to inadvertently code some controlled units as never-controlled, our data analysis could erroneously detect spillovers that reflect nothing more

[19] This database was constructed by double-entry hand-keying the four bound volumes of the 1995 Cambridge Assessor's Commitment Books, which were provided to us by the Cambridge Historical Commission.

[20] While we filed our own FOIA request with the City of Cambridge, we ultimately utilized the file obtained by David Sims through an earlier FOIA because its coverage appeared more complete.

than appreciation of formerly controlled units after decontrol. To be conservative, we code *all* units on a map-lot as rent controlled if *any* unit at that map-lot was controlled in 1994. It is therefore very unlikely that there are controlled units that we fail to capture. Conversely, when measuring the rent control *intensity* of a given geographic area, we calculate the fraction of residential units—rather than structures—that are rent controlled.[21] This is also conservative in that it prevents us from overestimating units' exposure to other controlled properties.

Figure 1 illustrates the prevalence of rent control in Cambridge, with dark circles indicating controlled properties. In 1994, 22 percent of all residential structures and 38 percent of residential units were subject to rent control. The dense neighborhoods close to the two major universities and proximate to the subway that bisects Cambridge from east to northwest contain high concentrations of renters and multi-unit structures and thus had relatively high rent control intensity. The largely owner-occupied area of Southwestern Cambridge features a higher fraction of single unit houses and hence had relatively low rent control intensity. It bears emphasis that our statistical analysis abstracts from these gross geographic differences in rent control intensity by comparing changes in residential prices among properties that differ in their proximity to controlled units but lie within relatively small neighborhoods.

We append two databases to analyze the impact of rent decontrol on market capitalization, the first enumerating property assessments and the second enumerating real estate transactions. The 1995 and 2005 Cambridge Assessor's files, which report property valuations from 1994 and 2004, provide the assessed appreciation of each extant property from the year prior to rent decontrol to nine years thereafter. The second is a commercial database provided by the Warren Group that enumerates all changes in ownership of residential properties for the years 1988 through 2005. Sourced from records of deeds, these data log each real estate transaction, including sale price, address, map-lot, number of bedrooms and bathrooms, lot size, year built, and property type. We exclude commercial properties such as apartment buildings from the analysis because such sales are rare and transact at heterogeneous prices that are in some cases extremely high.

Assessments and transactions provide complementary means to measure the capitalization of rent control's end. Assessments, our preferred measure, contain the universe of residential properties along with assessed market values at two points in time, immediately prior to rent control removal and ten years later. Assessments may offer a lagging indication of residents' changing willingness

[21]Our data always allow us to calculate the share of units in a building that are controlled, though we often cannot determine which specific units these are.

to pay for locations, however, and could differ from market valuations due to discretionary aspects of the assessment process. The sales data, in contrast, include both market prices and a rich set of property characteristics for locations where transactions take place, and, because they are available annually,provide a clearer picture of the trajectory of property price changes. Only a small percentage of residential units transact each year, however, and hence the sales data contain information on an incomplete and potentially non-representative set of residential units (we subsequently analyze whether rent control impacted the composition of transacted properties).

Table 2 presents descriptive statistics for the assessed Cambridge residential houses and condominiums used in our analysis, comprising 15,475 properties in 1994 and 17,505 in 2004.[22] Slightly more than half of these properties are houses. Rent controlled properties account for 29 percent of all residential properties, with condominiums comprising the substantial majority. Because the vast majority of Cambridge houses were and are owner-occupied, only 12 percent of houses were ever subject to rent controls.[23] House prices rise substantially in real terms during our sample period: the average 1994 assessed value of a decontrolled condominium is \$116,000, while it is \$351,000 in 2004—an increase of 111 log points.[24] Houses typically have higher assessed value than condominiums, and in both periods, decontrolled houses and condominiums have lower values on average than never-controlled houses.

2.3.2 Measuring rent control intensity (RCI)

Gauging each residential property's rent control exposure requires a metric that specifies which nearby units should be counted in the unit's reference set—that is, to which units it is 'exposed'—and how the rent control status of these reference units should be aggregated into an exposure index. For most analyses, we calculate the rent control status of the surrounding units to which a given property i is exposed by summing the number of controlled units within a surrounding geography g and dividing it by the sum of all residential units J_g (controlled and non-controlled)

[22]Note that a property may contain multiple units, e.g., a multi-family house.

[23]The house and condominium designations in Table 2 reflect the property's residential category at the time of assessment.

[24]Prices are deflated by the Consumer Price Index for All Urban Consumers, Series Id CUUR0000SA0L2. This index is an average for U.S. cities and excludes the price of shelter since we do not wish to confound the outcome measure, house price appreciation, with the numeraire.

in that geography:[25]

$$\text{RCI}_{i(g)} = \frac{1}{J_g} \times \sum_{j \neq i}^{J_g} \text{RC}_{j(g)}.$$

In a subsequent sensitivity analysis, we calculate each unit's rent control exposure as an exponentially declining function of its distance from all other controlled and never-controlled properties in the city.

The second input into the exposure measure is the choice of a surrounding geography. One potential set of geographies is supplied by the U.S. Census Bureau, which subdivides the area of cities into three increasingly fine geographic units: tracts, block groups and blocks, of which there are 30, 89, and 587, respectively, in Cambridge containing at least one assessed house or condominium.[26] While these pre-defined Census geographies have the virtue of allocating Cambridge land parcels into exhaustive, mutually exclusive geographic units, they have two substantial drawbacks for our analysis. One is that the Census geographies do not necessarily correspond to any specific notion of neighborhoods or proximity. For example, Census blocks frequently divide streets down the center, so that units on opposite sides are assigned to different blocks, which is clearly undesirable for measuring spillovers from nearby properties. The second is intrinsic to any allocation of geography into non-overlapping parcels: units closer to the perimeter of a geography are treated differently from units located in its center. For example, for a residential unit located on the northern edge of a geography, its neighbors 50 feet to its south will contribute to the unit's rent control exposure measure whereas its neighbors 50 feet to its north will not. By contrast, for a unit located in the center of a geography, its equidistant neighbors contribute equally to its rent control exposure measure.

To avoid both drawbacks of using fixed geographies, our preferred measure of a unit's rent control exposure is the fraction of residential units within a fixed straight line radius of 0.10, 0.20, and 0.30 miles of that unit that were controlled as of 1994. This radius exposure construct non-prejudicially selects the residential units that are physically closest to the reference unit [27] To provide a feel for

[25]Although our analysis of assessed values and transactions excludes apartment buildings, both controlled and never-controlled apartments contribute to the numerator and denominator of our exposure measure. Each rental unit within a multi-family house is counted separately in both the numerator and denominator. The RCI determination for a condominium structure excludes all other units in that structure.

[26]These units have average land areas of 0.22, 0.07, and 0.01 square miles respectively in Cambridge, housed an average of 3,145, 986, and 135 residents in 1990, and contained a mean of 1,292, 428, and 63 residential units in 1990. Additional details on the size, population, and number of structures and units in Census geographies are contained in Table A3.

[27]Because we calculate RCI using only Cambridge properties, the radius-based RCI for properties close to the City's edge excludes nearby units that lie outside the city. To address this source of potential mis-measurement, we

the area encompassed by these radii, Figure 1 plots concentric rings of appropriate scale overlaid on the Cambridge map.

Our main estimates are based on rent control intensity measured at a radius of 0.20 miles, which corresponds to about 0.13 square miles—an area larger than a block group but smaller than a tract in our sample. For the typical residential property, 34 percent of the surrounding units within a 0.20 radius are rent controlled. As shown in Table 2, condominiums are in neighborhoods with more rent control than houses, and both decontrolled houses and condominiums tend to be in more rent control intensive neighborhoods than their never-controlled counterparts. For instance, in 1994, 32 percent of units surrounding a typical never-controlled condominium are controlled, compared to 45 percent for decontrolled condominiums. There is also considerable cross-sectional variation in rent control intensity. Across all assessed properties, the standard deviation of RCI measured at 0.20 miles is 17 percentage points, and the range of the RCI measure spans from 0 to 72 percent.

2.4 Capitalized Effects of Rent Decontrol: Evidence from Assessments

Our illustrative model suggests that the capitalization of rent decontrol should accrue through three channels: the direct effect on decontrolled properties of the elimination of price controls and condominium conversion restrictions and associated investments; the indirect effect of decontrol on the desirability of neighborhoods in which controlled properties were located, stemming from improvements in neighborhood amenities—e.g., better upkeep, more desirable neighbors—and potentially affecting the market value of both decontrolled and never-controlled properties; and finally, broader increases in the desirability of Cambridge as a residential location, which may accrue city-wide.

Our econometric model recognizes each of these channels. We fit equations of the form:

$$\log(Y_{igt}^{A}) = \gamma_g + \delta_t + \beta' X_i + \lambda_1 \cdot \mathrm{RC}_i + \lambda_2 \cdot \mathrm{RCI}_i + \rho_1 \cdot \mathrm{RC}_i \times \mathrm{Post}_t + \rho_2 \cdot \mathrm{RCI}_i \times \mathrm{Post}_t + \epsilon_{igt}, \quad (2.2)$$

where Y_{igt}^{A} is the real assessed value of property i in neighborhood g in year t, γ_g are fixed effects representing different geographies, δ_t are year effects, and X_i are property characteristics such as

have verified that our findings are robust to discarding properties on all Cambridge block groups that directly abut the towns of Somerville, Arlington, Belmont, Watertown (except those that border the Charles River, the sizable Mt. Auburn Cemetery, or the light rail system in the southeast of Cambridge).

housing type (condominium, single family, two-family or three-family house). The dummy variable RC_i is equal to one for properties that were rent controlled in 1994 (prior to the law's repeal), while the Post indicator is equal to one for 2004. Of central importance to the analysis, the variable RCI_i measures the fraction of units nearby to i that were controlled as of 1994. Our main specifications code "nearby" units as those within a 0.20 mile radius of a given property, but we subsequently explore alternative definitions. Recognizing that real estate prices of nearby properties are not independent, we generally cluster the standard errors at the level of Cambridge block groups.

The coefficient ρ_1 estimates the direct effect of rent control removal on the assessed value of formerly controlled properties by contrasting the change in value of controlled versus never-controlled properties following the end of rent control, holding constant unit characteristics, cross-neighborhood differences in residential real estate prices and over-time, and city-wide changes in residential real estate prices. The coefficient ρ_2 estimates the indirect effect of rent decontrol on the value of decontrolled and never-controlled properties by contrasting changes in the value of units in geographies with high rent control intensity relative to those with low rent control intensity, again holding constant property characteristics, neighborhood effects, and time effects. Finally, any effects of decontrol that accrue city-wide—that is, are not limited to decontrolled properties or the neighborhoods in which they were located—are absorbed by the time effects δ_t. Since these time effects soak up any macroeconomic factor affecting the value of Cambridge's housing stock in this time period, we do not interpret the evolution of δ_t as a causal effect of rent decontrol.[28]

For ρ_1 and ρ_2 to provide unbiased estimates of the direct and indirect effects of rent decontrol on the market value of residential properties, it must be the case that the elimination of rent controls—and resulting neighborhood level changes—must not have been fully anticipated by households and landlords. This appears plausible in light of the fact that the rent control law was narrowly eliminated (51 to 49 percent) by a state-wide referendum in which a large majority of Cambridge residents voted against rent decontrol.[29] Additionally, our identification requires that conditional on detailed geographic and time effects, the variable representing a property's exposure to rent decontrol ($RCI_i \times Post_t$) is uncorrelated with other unmeasured factors within neighborhoods that

[28] We hesitate to interpret the coefficients on the RC main effect and RCI $\times$ RC (coefficients λ_1 and λ_2) as causal effects of rent control status or rent control intensity since these variables will also pick up unobserved factors that determined rent control status and rent control intensity at the time that rent control was adopted in 1970 (for example, the age of the residential housing stock and the fraction of nearby units that were owner occupied versus rented).

[29] To the degree that rent decontrol (and any resulting neighborhood effects) were foreseen by incumbent and potential owners, buyers and renters, these effects would substantially capitalize into values before rent control was removed, which would work against our finding either a direct or indirect effect of rent decontrol on prices.

affect local house prices, change contemporaneously with rent control removal, but yet are not caused by the elimination of rent control. It is difficult to state precisely what these factors would be since the most obvious candidates (e.g., improvements in neighborhoods) are plausibly caused by rent control removal. We subsequently present event-study graphs with the transaction price sample that strongly suggest that the effect of rent control intensity on house prices is not present prior to the elimination of rent control and evident thereafter. In many instances, we also estimate a richer version of equation (2.2) in which we interact the RCI measure with both the rent control main effect and the RC $\times$ Post term. This triple-difference specification allows the indirect effect of rent control intensity to differ between controlled and never-controlled properties in both the rent control and decontrol eras.

The end of rent control in 1995 coincided with a period of nationwide house price appreciation, which raises the possibility of confounding price trends. The time effects δ_t in our estimating model will absorb these changes to the degree that they affect the overall price level of Cambridge housing. They will not absorb any differential appreciation in rent control-intensive neighborhoods, which might hypothetically occur if, for example, the U.S. housing boom of the early 2000s spurred an influx of lending (and associated price appreciation) in rent control-intensive neighborhoods (Mian and Sufi, 2009). We address this concern by estimating specifications containing tract-by-year interactions, in addition to 89 geographic main effects for Cambridge block groups, thereby allowing the rate of appreciation to differ across Census tracts.

2.4.1 Appreciation of decontrolled properties

Table 3 presents baseline estimates of equation (2.2) for the causal effect of rent decontrol on assessed values of decontrolled properties from 1994 to 2004 using the full set of 15,475 residential properties. Column 1 reports a parsimonious specification containing only an RC main effect, an RC$\times$Post indicator, and a set of dummies for year-of-sale and structure type (condominium, two family house, three family house). Prior to rent decontrol, the assessed value of controlled (RC) properties averaged 50 log points below the assessed value of never-controlled (non-RC) properties.[30]

[30]The RC main effect estimates do not admit a causal interpretation, as noted above. A property's rent control status in 1994 is a function of the property's year of construction and its residential and occupancy status (rental vs. owner-occupied) as of 1971, which in turn are likely to be correlated with the fixed characteristics of the property, its maintenance and appearance, as well as the desirability of its surrounding neighborhood. While the rent control main effect is robustly large and negative in all cases, this may reflect omitted property attributes and not the causal impact of rent control.

Following decontrol, this gap closed by 22 log points. Columns 2 through 4 refine the precision of the comparison by adding a set of dummy variables that sweep out cross-neighborhood differences in price levels and trends. Column 2 adds block group effects, which absorb average assessed values within narrow block groups (averaging 0.07 square miles). Here, the model is identified by contrasting the change market value of decontrolled and never-controlled units within block groups. Column 3 adds tract by year dummies, thus allowing each of the 30 Census tracts in Cambridge (averaging 0.22 square miles in area) to have a different overall appreciation rate. The final column includes a fixed effect for each residential location or map-lot (a total of 9,497 map-lots). This demanding specification, which absorbs the RC main effect, contrasts the map-lot level change in assessed values between map-lots that contained controlled units versus those that did not, again allowing for different price trends across 30 Census tracts.

Across all specifications, the rent control main effect is highly robust and highly stable, demonstrating that decontrolled units appreciated substantially relative to never-controlled properties. These initial estimates do not distinguish, however, between the direct and indirect channels that may jointly contribute to this appreciation. In particular, since decontrolled properties are typically located in neighborhoods with above average levels of rent control intensity (Table 2), the RC$\times$Post term estimated above captures a combination of direct decontrol effects and indirect (micro-neighborhood level) effects stemming from the greater desirability of formerly rent control-intensive locations. Our next set of estimates distinguishes these two effects.

2.4.2 Direct and indirect effects of rent decontrol

Table 4 augments the simple difference-in-difference models above with a measure of the rent control exposure of each residential property (denoted RCI and calculated using a 0.20 mile radius), as well as an interaction term between the RCI measure and a post-1994 indicator variable. This term measures the degree to which properties with greater rent control exposure saw differential appreciation following decontrol. The inclusion of the RCI and RCI $\times$ Post measures also changes the interpretation of the RCI $\times$ Post main effect. Whereas previously this variable measured the differential appreciation of decontrolled versus never-controlled properties averaging (implicitly) across more and less rent control-intensive areas, the RCI$\times$Post coefficient in the augmented specification measures the differential appreciation of RC relative to non-RC properties in a hypothetical location with no other surrounding controlled properties (that is, RCI = 0).

The base specification in column 1, which contains only year of sale and structure type dummies in addition to the RC and RCI terms, finds that properties with higher rent control exposure had lower value in the decontrol era, and that this differential was substantially reduced in the period following decontrol. Specifically, the point estimate of -0.58 on the RCI measure indicates that a property at the mean level of rent control exposure of 0.32 was assessed at approximately 19 log points below a property with zero exposure. We do not take the main effect of the RCI variable to be causal, however, since it is likely to be correlated with the many factors that determined which properties were controlled in 1971. Conversely, the coefficient of 0.33 on the RCI $\times$ Post indicator implies that 56 percent of this price differential was erased in the years after decontrol. Under our identifying assumption that these unobserved factors are quasi-fixed or are not spuriously correlated with rent control intensity across local areas, the RCI $\times$ Post interaction may be viewed as a causal estimate of the indirect effects of rent control on the market value of surrounding units (both formerly and never-controlled). The fact that both the RC main effect and the RCI $\times$ Post coefficients fall in magnitude relative to the Table 3 estimates (which exclude the RCI measure) reveals that the lower market value of RC properties stems in part from the fact that they were situated in more rent control-intensive locations.

We explore the robustness of these initial relationships by applying the control variables used above: block group fixed effects, tract-year effects, and map-lot fixed effects. While these covariates reduce the precision of the RCI main effect, the point estimates remain large and statistically significant. The coefficient of primary interest (RCI $\times$ Post) increases in magnitude with the inclusion of tract-year effects. Column 3 obtains an estimate for the RCI $\times$ Post coefficient of approximately 55 log points, while in column 4 the estimate is 48 log points, implying that a residential property at the 75^{th} percentile of rent control exposure gained approximately 13 percent more in assessed value following decontrol than a property at the 25^{th} percentile of exposure.[31]

These first four estimates constrain the indirect effects of rent control to be identical for never-controlled and decontrolled units. In practice, these effects may differ. Indeed, if the indirect effect were only present for decontrolled units, this would suggest that the indirect effect is not operating through the hypothesized localized amenity channel (which we would expect to affect both property types). The models in columns 5 through 7 demonstrate that both decontrolled and never-controlled properties benefit from the indirect effect of rent control removal. In the

[31] The 25th and 75th percentiles of the RCI distribution are 0.464 and 0.199. The implied interquartile effect is $0.126 = 0.475 \times (0.464 - 0.199)$.

most demanding specification in column 7, which includes map-lot fixed effects and tract by year dummies, we estimate an RCI $\times$ Post coefficient of 0.42 for never-controlled properties and 0.61 for decontrolled properties. Both are significantly different from zero, and the data do not reject the hypothesis that these coefficients are of the same magnitude.

Notably, the models that allow for separate indirect effects for controlled and never-controlled properties also find that the RC discount is only approximately half as large as was implied by the earlier models that do not include interaction terms between RCI and rent control status—approximately 23 rather than 48 log points—and that this RC discount was fully offset by the post-decontrol appreciation of decontrolled properties. By implication, approximately half of the estimated RC discount detected in columns 1 through 3 of the table is accounted for by the fact that RC units were situated in more rent control-intensive locations, and that rent control exposure differentially lowered their value. While our conceptual model is silent on why the indirect effect of rent control is greater for controlled than never-controlled properties, one speculative explanation is that deferred maintenance and poor property management were more acute in locations where a larger fraction of properties was controlled. This conjecture would also be consistent with our finding of greater relative appreciation of decontrolled than never-controlled units in rent control-intensive locations– though as above, this differential is not statistically significant.

2.4.3 Variation across property types

Table 5 explores the potentially differing consequences of rent decontrol for the assessed values of houses and condominiums.[32] Across the two panels, the direct impact of rent decontrol on controlled houses is substantially smaller than the corresponding estimate for condominiums, a pattern that may be due to the greater extent of upgrading at controlled condominiums.[33] For residential houses (upper panel), we estimate an indirect effect coefficient on the value of residential houses of approximately 20 log points, implying that a house facing the mean level of RCI of 0.37 would experience an additional 7.4 log points of appreciation relative to a non-exposed house following rent decontrol. When allowing for separate indirect effects for decontrolled and never-controlled houses (columns 3 and 4), we find that the indirect effect for decontrolled houses is 20 to 50 percent

[32]To simplify exposition, we display only the interaction terms between post-decontrol and the RC and RCI terms, suppressing the included main effects of these variables.

[33]As discussed in section 2.7, Cambridge building permit data indicate that annual city-wide investments in decontrolled condominiums increased by 206 percent in the post versus pre-decontrol period while the corresponding increase for decontrolled houses was 120 percent (Table A1).

larger than for never-controlled houses, although we are unable to reject the equality of the two regression coefficients. Adding 9,497 map-lot fixed effects to the regression model in column 4 decreases precision such that the indirect effect estimates for houses become insignificant, though magnitudes are only modestly affected.

The parallel analysis for condominiums in the lower panel finds significant indirect effects of rent decontrol on both decontrolled and never-controlled condominiums. The indirect effects for condominiums are greater than for houses, although standard errors are also considerably larger. The point estimate of 0.49 in column 2 implies differential appreciation of 18.2 log points for a condominium at the mean level of rent control exposure relative to a non-exposed unit. As is the case with houses, adding a map-lot fixed effect for each land parcel containing condominiums (1,450 fixed effects) reduces or eliminates the statistical significance of the indirect effect point estimates.

The number of condominiums in Cambridge rose by one third between 1994 and 2004, with almost half of this rise due to the conversion of existing houses to condominiums. This substantial change in the housing stock implies that part of the rise in assessed values may be due to capital improvements in residential units, particularly condominiums, rather than solely changes in the value of ownership stemming from decontrol (e.g., the option to charge higher rents or convert the unit to owner-occupied status). Notably, this concern applies only to the *direct* effects estimates of rent decontrol on decontrolled units, which may conceivably combine both the investment and ownership channels. For the indirect effects we measure, this source of variation—spillovers from local housing investments spurred by rent decontrol—is not a concern; indeed, this is one of the key causal channels through which we hypothesize the indirect effects operate.[34]

To explore the importance of the investment channel, we undertake two exercises. We first re-estimate the main models for direct and indirect effects of decontrol on houses and condominiums while excluding all units that changed usage categories, e.g. converted condominiums, between 1994 and 2004. Then, in section 2.6, we directly explore the role of investments using Cambridge residential building permit data.

When dropping converted properties from the regression estimates in the final two columns of Table 5, we find little effect on the direct or indirect effect point estimates for houses. This is sensible since

[34]The converse concern applies, however: if localized spillovers spur additional investments at decontrolled units, this may contribute to the estimated *direct* and indirect effect, meaning that our interpretation of the direct effect estimate is unduly restrictive. However, to the degree that these spillover-induced investments are greater in more rent control-intensive locations, as we anticipate, the indirect effect coefficient should correctly capture this channel.

only one in ten houses changed status during the ten-year window. Consistent with our reasoning above, however, the estimated *direct* effect of rent decontrol on condominium values is substantially reduced when converted properties are excluded—falling by as much as 35 percent between columns 3 and 5—while the indirect effects estimates are substantively unaffected.[35] These results suggest that the direct effects estimates capture both capital improvements and changes in ownership value spurred by decontrol, particularly for condominiums. The indirect effects estimates, however, are not affected (conceptually or empirically) by abstracting from the substantial investments made in converted units.

2.4.4 Testing alternative measures of rent control intensity

Our estimates so far employ a measure of rent control intensity calculated over a 0.20 mile radius from each Cambridge map-lot. We explore the sensitivity of the results to this choice by employing two sets of alternative measures, one that varies the geography over which rent control intensity is calculated, and one that varies how the weight given to surrounding properties decays with distance. These robustness tests employ the final (most exhaustive) specification in Table 4, which includes fixed effects for each individual map-lot and tract-by-year dummies that account for differing rates of property appreciation across all thirty Cambridge Census tracts.

The upper panel of Table 6 reports estimates using RCI measures calculated at differing geographies. The first three columns calculate RCI at radii of 0.10, 0.20 and 0.30 miles, respectively (hence, column 2 replicates our main specification from Table 4). Column 4 instead uses an RCI measure calculated at the level of 587 Census blocks. Distinct from the radius-based measures, Census blocks comprise a set of contiguous, non-overlapping geographic subdivisions. These blocks may not, however, correspond to any specific notion of neighborhood or proximity, particularly since residential units on opposite sides of the same street are often assigned to different blocks.

The direct impact of decontrol is relatively insensitive to the geography of the RCI measure. Across specifications, this effect averages 13 to 15 log points, which is generally not significant.[36] The indirect effect estimates are somewhat more sensitive. The largest indirect effect estimate comes from the 0.30 mile radial RCI measure, though this effect is statistically indistinguishable from

[35] This is logical since excluding converted units from the regression should not eliminate the *indirect* effect of investments made at these units on the value of non-converted units.

[36] As shown in Table 4, the RC main effect is generally a much smaller and less precise in models that include three way interactions between the RC, RCI and Post variables, reflecting the fact that prior to rent decontrol, controlled units were valued at the greatest discounts in more rent control-intensive neighborhoods.

the 0.20 mile measure. Estimates using the Census block group based RCI measure are generally insignificant, which is consistent with our observation that block boundaries do a poor job of capturing neighborhood proximity.

An important conceptual limitation of the radius-based measure is that it puts equal weight on each residential unit within a specified radius of a given map-lot, while simultaneously according zero weight to all other units in Cambridge. It seems plausible, however, that the interactions among residential units decline with distance, so that nearby units matter more for a unit's rent control intensity as perceived by occupants and potential buyers, while more distant units matter less.

To explore this idea empirically, we employ an alternative measure of rent control intensity that places greatest weight on nearby units and less weight on more distant units. Specifically, we use an exponential decay function to calculate the RCI of each unit i as a function of its distance from all other RC (rent controlled) units $j \neq i$ in Cambridge, where the weight given to each unit j is declining in its distance from i. Let d_{ij} be the distance between units i and j measured in miles and $\lambda < 0$ be a negative constant, J be the complete set of residential units in Cambridge, and RC_j be a dummy variable equal to 1 if unit j is rent controlled and 0 otherwise. Our distance-based measure of RCI^λ is:

$$\mathrm{RCI}_i^\lambda = \frac{\sum_{j \neq i}^{J} \mathrm{RC}_j \times e^{\lambda d_{ij}}}{\sum_{j \neq i}^{J} e^{\lambda d_{ij}}}. \tag{2.3}$$

Like the primary RCI measure, the measure RCI_i^λ lies on the unit interval. The difference between RCI_i^λ and RCI is that the weight given to surrounding units in RCI_i^λ is a continuous, declining function of distance from i whereas for RCI, the weighting function is flat over the area of the designated radius, and then is equal to zero outside of that area.

The lower panel of Table 6 reports estimates of this decay-based RCI measure using values of λ ranging from -12 to -3, where lower (more negative) values of λ give greater weight to nearby units and higher values of λ give greater weight to more distant units. To illustrate the operation of our weighting function, the first several rows of the panel display the weight accorded to units at 0.10, 0.20, and 0.30 miles from each reference unit relative to units at 0.01 in the RCI calculation. All four values of λ accord a weight that is close to unity to properties within a radius of 0.01 miles and under, whereas more negative values of λ places substantially less weight on distant units.[37] For example, at $\lambda = -12$, units at 0.20 and 0.30 miles receive weights of 0.10 and 0.03 relative to the weight at 0.01 miles, respectively. In contrast, with $\lambda = -3$, these units receive relative

[37] Units at distance zero always receive a weight of one since $e^0 = 1$.

weights of 0.57 and 0.42, which are substantially greater. The estimated direct and indirect effects of rent decontrol on residential property appreciation are both stable and robust across the four parameterizations of the decay function. In all cases, the indirect effects estimates are statistically significant for both decontrolled and never-controlled properties, and are comparable in magnitude to the radius-based measures –though if anything, the decay-based estimates are more robust.

One pattern evident in both panels of Table 6 is that the estimated magnitude of the indirect effect rises when we use an RCI measure that gives greater weight to more distant properties (by employing a wider radius or a more gradual decay function). A likely explanation for this pattern is that employing a broader RCI measure provides more information about the extent of a unit's RCI exposure; that is, a given high (or low) RCI value obtained over a larger radius (or slower decay function) implies that the relevant unit is more (or less) deeply surrounded by other controlled units. This should in turn imply a larger indirect effect of RCI on market values. Logically, the variance of the RCI measure declines as its scope broadens as shown in the lower row of each panel, so the size of the standardized effect rises less rapidly than do the RCI point estimates.[38]

Alongside the robustness tests in Table 6, we have in the online appendix explored a variety of alternative and complementary identification strategies that probe the key results. One potential limitation of our primary approach stems from the non-parallelism between the RCI measures and the geographic dummy variables used as controls. While the radius-based RCI measure in many cases partially overlaps multiple block groups, each map-lot in the regression is associated with only a single block group fixed effect. To explore sensitivity to this choice, we created a set of "rolling" block fixed effects. For each map-lot, we identify each Census block (of which there are 587) whose centroid lies within 0.2 miles of the map-lot and assign these block dummies to the map-lot. These non-mutually exclusive block fixed effects are then used in place of the conventional block group fixed effects in the regressions (see Table B1). We perform an analogous exercise for the exponential decay specifications, where all block dummies in Cambridge are fractionally assigned (summing to unity) to each map-lot as a decaying function of the distance between the map-lot and block centroids (see Table B1). Motivated by the fact that our radius-based rent control intensity cannot account for the characteristics of non-Cambridge properties bordering the city, we perform a third specification test that obviates this issue by excluding all Cambridge block groups that border non-Cambridge properties (see Table B2). All three of these sensitivity exercises, reported

[38]At maximal radius, the RCI measure is identical for all Cambridge units save for each map-lot's own effect on the RCI measure (since units are not counted in their own RCI measures).

in the online appendix, yield estimates that are highly comparable in magnitude and precision to our primary estimates above.

2.5 The Time Path of Rent Decontrol Capitalization

Because our assessor data covers only two points in time, 1994 and 2004, they do not shed light on how residential real estate prices evolved prior to decontrol or in the years thereafter. We turn to housing transaction data to complete this picture. We begin with simple "event study" plots of the main effect of rent control status on real estate transactions, estimated with the equation

$$\log(Y^S_{igt}) = \gamma_g + \delta_t + \beta' X_i + \sum_{t=1988}^{2005} (\mathrm{RC}_i \times \delta_t)\rho_{1,t} + \epsilon_{igt}, \tag{2.4}$$

where Y^S_{igt} is the real sales price of residential unit i located in block group g in year t, the vectors γ_g and δ_t contain fixed effects for block groups and year of sale, and the X vector contains a rich set of property characteristics, sourced from deed records and summarized in Table A4, including the count of rooms, bathrooms, and bedrooms, the unit's interior square footage, a quadratic in lot size and a dummy for lot size equal to zero (commonplace for condominiums), and a quadratic in the log of property age and a dummy for missing year built. All controls are interacted with dummies for structure type (condominium, single family home, multi-family home) since the hedonic value of these attributes may differ across types, and structure type dummies are further interacted with quadratic time trends to allow for differing price trends. Standard errors are clustered at the block group level. Figure 3 plots the key coefficients ($\rho_{1,t}$) from equation (2.4), which correspond to by-year estimates of the rent control price differential measured relative to the omitted reference year of 1994.

The relative price of RC properties increased by roughly 10 log points over the first three years following decontrol, declined very modestly between years three and four, and then rose almost continuously thereafter. By the end of the sample in 2005, RC properties had increased in market value by almost 30 log points relative to nearby non-RC properties with similar characteristics. The increasing cumulative effect of decontrol on transaction prices parallels the evidence above on the evolution of resident turnover, which also rose immediately following decontrol and then generally trended upward through the end of the sample window. Both results suggest that changes in the desirability of locations and neighborhoods induced by decontrol likely took years to unfold.

We plot the indirect effects of rent decontrol on the value of never-controlled and decontrolled properties in Figure 4 using a specification analogous to equation (2.4) augmented with RCI×Year terms.[39] Indirect price effects of decontrol on the sale prices of never-controlled properties begin to accumulate immediately following decontrol, attain statistical significance by the fifth year following decontrol (1999), and continue to rise through the end of the sample, yielding a point estimate of 0.60 log points in 2005 ($p < 0.01$). Indirect effects estimates for decontrolled properties (panel II) offer a similar picture of post-decontrol appreciation: the indirect effect for decontrolled units varies in sign and is generally insignificant in the pre-decontrol years; following decontrol, the point estimates are strongly positive in 10 of 11 years and are statistically significant at $p < 0.10$ in 7 of 11 years. Though the event study plot is substantially noisier for decontrolled than for never-controlled properties, this likely reflects the fact that there are only half as many transactions of decontrolled units between 1988 and 2005 (4,802 relative to 9,987).

Table 7 explores these relationships in further detail. The first three columns provide estimates of the direct effect of decontrol on the market value of formerly controlled properties. The estimates range from 6 to 11 log points as we add block group fixed effects and controls for property characteristics (column 2).[40] Taking advantage of the additional years of data available from the transactions sample, column 3 adds linear and quadratic trends for each of the 30 Census tracts (also interacted with property types) to allow for the flexible evolution of real estate prices over time within fine geographies. These controls have little impact on magnitudes or precision. Subsequent columns introduce the RCI measure to estimate indirect effects. The coefficient of interest in column 4 (RCI×Post) is 21 log points, which is in the lower range of estimates obtained using the assessor's sample (Table 4). As Figure 3 suggests, however, both the direct and indirect effects of rent decontrol cumulate over the sample window, so that the average post-decontrol effect is smaller than the long run effect estimated by the long difference specifications used with the Assessor data. Adding flexible geographic trends (column 5) slightly reduces the indirect effect estimate and increases its standard error.

Columns 6 and 7 report estimates where the indirect effect of rent control exposure on prices is permitted to differ between decontrolled and never-controlled properties. The indirect effect estimates

[39] The specification is estimated separately for never-controlled and decontrolled units and hence RC x Year effects are not included.

[40] That the addition of property-specific covariates in column 2 substantially reduces the magnitude of the baseline price differential between RC and non-RC properties (from 30 to 20 log points) is consistent with the data summarized in Table A4, which indicates that RC properties are situated on smaller lots and in older structures, and in the case of condominiums, provide less square footage than non-RC properties.

are statistically indistinguishable between decontrolled and never-controlled units, though they are generally larger for decontrolled properties and more precisely estimated for never-controlled properties (consistent with Figure 4). In the final column, which includes all prior covariates plus tract-specific quadratic trends, the point estimates for the indirect effects are no longer significant at conventional levels, though the point estimates are little affected.

Echoing the results from property assessments, the transaction data models suggest that the price penalty for rent control exposure was substantially greater for controlled than never-controlled houses in the rent control era. Accounting for this differential in columns 5 through 7 reduces the magnitude of the estimated RC price discount prior to decontrol. For example, the RC point estimate of -17 log points in column 6 is quite comparable to the RC point estimates in Table 4 that use property assessments and allow for separate RCI slopes for RC and non-RC units. While the direct effect of decontrol on RC units appears modestly smaller in magnitude in the transaction-based models than in the assessment-based models, this again likely reflects timing, with the assessment-based models using a 10-year long change and the transaction-based models using a series of annual observations.

The transaction sample also provides an opportunity to explore the concern that assessed real estate values may not accurately reflect market prices. We examined this issue by constructing a matched assessor's sample for properties that transacted in 1994 and 2004, and we used this matched set to perform parallel estimates of the direct and indirect effects of decontrol on assessed and transacted property values in Table A5. These models yield a close match between the estimated decontrol impacts on assessed values and transaction prices and, moreover, the match is particularly close for houses.[41] This comparison suggests that the transaction and the assessor's sample provide complementary and broadly consistent measures of valuations.

We also explored the possibility that rent decontrol had positive indirect effects on the market value of properties close by Cambridge, focusing on the adjoining town of Somerville, Massachusetts. For this exercise, we assembled residential transaction records for Somerville analogous to those used above for Cambridge, limiting the sample to residential units located in Census tracts and block groups that abut either the Somerville-Cambridge border (North of Cambridge) or the Somerville-

[41] An additional complexity in comparing the assessed versus sale values of condominiums is that we are unable to determine which specific unit among the assessed condominiums units at a map-lot is transacted. Consequently, we include matched assessor data for *all* units at the map-lot where one or more unit transacts. This leads to a sample of 7,897 condominium assessments matched to the 937 units that were transacted in 1994 or 2004.

Medford border (North of Somerville).[42] Comparing properties along these two Somerville borders, in Table A6 we find robust evidence that Cambridge-bordering properties appreciated by 7 to 10 percent more than did Medford-bordering properties with the same observable characteristics. This pattern is consistent with the hypothesis that improvements in Cambridge neighborhood amenities following decontrol also increased the desirability of locations bordering Cambridge.

The online appendix reports on additional robustness tests that investigate price appreciation in neighboring towns, explore the potential importance of the increase in subprime credit in lower-income Cambridge neighborhoods during this time period, and test for correlations between rent control-intensity and changes in the composition of Cambridge properties that transact. These many specification tests, in combination with our prior estimates using assessed property values, confirm that alongside its direct effects on the market value of formerly controlled units, rent decontrol had robust indirect impacts on market values. Both decontrolled and never-controlled properties in Cambridge that were more exposed to controlled units saw differential appreciation in the post-decontrol regime. Before benchmarking the economic magnitudes of these direct and indirect effects, we briefly consider one plausible channel through which they may have operated: property investments.

2.6 The Impact of Rent Decontrol on Property Investments

Cambridge experienced an overall investment boom after the end of rent control. Total permitted investment at houses and condominiums rose from $83 million in the period 1991-1994 to $455 million in the period 1995-2004, while annual investment expenditures roughly doubled at three of four property types—decontrolled houses, never-controlled houses, and never-controlled condominiums—and roughly tripled at decontrolled condominiums (Table A1).[43] While fewer than one in twenty-five residential units receives a building permit annually, this fraction increased substantially following decontrol—by 17 and 7 percent among never-controlled and decontrolled houses,

[42]Somerville is bordered by Cambridge, Arlington, Medford, Everett, and Charlestown. Its two longest borders by a considerable stretch are those with Cambridge in the South and Medford in the North. The transactions data we assemble are also sourced from the Warren Group files, used for the price analysis immediately above, and contain the identical data elements and years of coverage.

[43]As detailed in the Data Appendix, our analysis in this section draws on a database of all building permits issued by the Cambridge Inspectional Services Department for years 1991 through 2005, including property address and proposed expenditure. Since permits can be filed either for a structure (e.g., a multi-unit condominium complex) or for any unit in a structure, we attribute a permit at a given structure to only one unit in that structure when computing permitted units in Table A1.

and by 38 percent and 45 percent among never-controlled and decontrolled condominiums.[44] Was this increase in residential investment caused by rent decontrol?

An event study of the direct impact of decontrol on permitting activity (Figure 5) shows a sharp, statistically significant differential rise in permitting activity and investments at decontrolled relative to never-controlled properties during the first five years following decontrol.[45] Relative to never-controlled units, the annual permitting probability at decontrolled units rose by 2.5 to 3.0 percentage points while permitted investment expenditures per decontrolled unit rose by \$1,000-\$1,500 dollars. We find no evidence, however, that decontrol indirectly caused permitting or investments to rise in decontrolled or never-controlled units with higher rent control exposure (not shown). Thus, any indirect effect of decontrol on residential investments that took place was not localized to highly rent control-exposed units—though it remains plausible that decontrol helped to spur the city-wide investment boom documented in Table A1. One factor that may obscure any direct expenditure effect in our analysis is low statistical power: the vast majority of investment expenditures are zero, while the mean and variance of expenditures at permitted units are high and rising (Table A1, Panel III). However, unless increased investment occurred along dimensions that do not require permits and hence are not observed in our data—e.g., repairs and maintenance that are not structural and do not alter major systems—the pattern of results appears to rule out very large differential expenditure effects at formerly controlled units. We consider an upper bound on the contribution of induced investments to post decontrol property appreciation in the next section.

2.7 The Capitalized Value of Rent Decontrol in Cambridge

How economically large are the direct and indirect effects of rent control estimated above? We answer this question by benchmarking our estimates against the overall level of house price appreciation in Cambridge using the Cambridge Assessor's Database as our measure of the value of the housing stock. Panel I of Table 8 presents information on the assessed value of the Cambridge housing stock in 1994 and 2004. Between 1994 and 2004, the assessed value of houses and con-

[44]We cannot exclude the possibility that the incentives to file for investment permits, or to accurately report investment costs on building permits, were affected by the rent control regime; for example, a landlord of a controlled unit might have been more likely to declare investment activity to justify a price increase to the Rent Control Board.

[45]By matching the permit data to the structures file, we observe permitting activity at every structure and hence our investment analysis sample is a balanced panel of structures by year, though the majority of permitting observations are zeros. We regress investment expenditure on a rent control indicator and the rent control indicator interacted with various post measures, controlling for year of sale dummies, and the number of units and its square (for multi-unit structures), structure type and structure times post interactions.

dominiums rose from \$4.7 billion to \$12.5 billion in constant 2008 dollars, a gain of 163 percent. Notably, appreciation of decontrolled units exceeded that of never-controlled units by a substantial margin: 219 percent relative to 152 percent, and these gains were larger among both decontrolled houses (185 percent versus 147 percent) and decontrolled condominiums (237 percent versus 166 percent).

What was the contribution of rent decontrol to these gains? We compute the direct and indirect contributions of rent decontrol to these valuations by applying our most conservative regression estimate, which includes map-lot fixed effects and Census tract trends (Table 4, column 7), to calculate the counterfactual price change at each location between 1994 and 2004, assuming that rent control had remained in place. While the aggregate value of decontrolled houses and condominiums increased from \$785 million to \$2.5 billion between 1994 and 2004, our estimates imply that had rent control not been eliminated, this gain would have been \$849 million smaller, with \$310 million due to the (foregone) direct effect of decontrol and an additional \$539 million due to the (foregone) indirect effects on decontrolled properties. By implication, slightly more than half of the \$1.7 billion appreciation of decontrolled properties between 1994 and 2004 can be accounted for by rent decontrol.

Though never-controlled units do not (by definition) benefit directly from decontrol, the indirect effects of decontrol are substantial. Our estimates imply that of the \$6.0 billion gain in the assessed value of never-controlled houses and condominiums between 1994 and 2004, \$1.1 billion (13 percent) of this gain is due to the indirect effects of rent decontrol on never-controlled properties, with \$822 million accruing to houses and \$306 million to condominiums.

Putting these components together, we calculate that decontrol added almost exactly \$2.0 billion to the value of the Cambridge housing stock between 1994 and 2004, with 84 percent of this effect due to the indirect effect of rent decontrol. While the share of post-decontrol appreciation between 1994 and 2004 induced by rent control removal was substantially larger for decontrolled than never-controlled properties (49 versus 19 percent), the never-controlled segment of the market received the largest increase in capitalization from rent control's removal: \$1.1 billion versus \$849 million. By implication, prior to decontrol, the never-controlled sector bore more than half of the incidence of rent control regulation.

Can the increase in residential investments documented in section 2.6 account for these price impacts? Total permitted residential investments averaged \$45.5 million between 1995 and 2004

(Table A1).[46] In the four years prior to decontrol, these expenditures averaged \$20.8 million. To benchmark the maximal estimate for the capitalized value of these investments, consider a case where the entire \$24.7 million in annual expenditures could be causally attributed to rent decontrol, where each dollar of expenditure led to a dollar of price appreciation, and where there was no subsequent depreciation of these investments during this ten year interval. In this case, we would conclude that only 12 percent of the appreciation of Cambridge residential properties between 1994 and 2004 was due to increased investments induced by rent decontrol (\$247 million of \$1,977 billion), leaving the remaining 88 percent accounted for by the capitalization of other benefits of rent decontrol. A parallel set of calculations implies that increased investments can explain at most 18 percent of the indirect effect of decontrol on the value of never-controlled properties, and at most 6 percent of the total (direct + indirect) effect of decontrol on the value of formerly controlled properties. In fact, our event study estimates in Figure 5 imply that decontrol can account for no more than \$82 million of the total increase in Cambridge residential investments in the years following decontrol, all of it concentrated on decontrolled properties.[47]

2.8 Conclusion

The largely unanticipated elimination of rent control in Cambridge, Massachusetts in 1995 affords a unique opportunity to identify spillovers in residential housing markets. This paper exploits the sharp cross-neighborhood contrasts in the fraction of units that were decontrolled to assess the localized price spillovers to never-controlled properties as well as to quantify direct effects on decontrolled properties. Our main finding is a large and significant positive indirect effect of decontrol on the valuation of properties that were exposed to controlled units, leading on average to a 16 percent increase in the value of residential units between 1994 and 2004. We further document that rent controlled properties were valued at a substantial discount relative to never-controlled properties, and that rent decontrol eliminated a substantial part of this differential, raising the assessed values of these properties by approximately 13 to 25 percent.

[46] The permitted investment cost may somewhat understate the full economic costs of housing investments since the property owner may also invest considerable oversight time on property improvements, and these costs are not included in building permits. Moreover, if the property owner must tie up working capital (or equivalently, obtain a construction loan) while improvements are made, then the full economic cost of investments should also include the opportunity cost of capital during the interval between investment and realized returns.

[47] For this calculation, we assume that decontrol caused each of the 5,439 decontrolled units extant as of 2004 (Table 2) to receive on average of \$1,500 per year of additional investment between 1996 and 2001 (Figure 6). This places an upper bound on additional induced expenditures during these years of \$49 million. Accounting for the 95 percent confidence interval surrounding this estimate, this number is \$82 million.

The contribution of decontrol to the capitalized value of the Cambridge residential housing stock in this period corresponds to a total of \$2.0 billion. While the direct effects on decontrolled properties were larger in percentage terms than the effects on never-controlled properties, the stock of controlled properties was smaller and less valuable than the never-controlled stock. As a consequence, indirect effects on never-controlled properties account for more than half (56 percent) of the decontrol-induced increase in the value of the housing stock.

Because under any reasonable set of assumptions, increases in residential investment stimulated by rent decontrol can explain only a small fraction of these indirect effects, we conclude that decontrol led to changes in the attributes of Cambridge residents and the production of other localized amenities that made Cambridge a more desirable place to live. This possibility is also highlighted by our theoretical model, though we are not able to thoroughly examine it with our data. Glaeser and Luttmer (2003) argue that non-price rationing under rent control leads to a mismatch between renters and apartments and provide evidence that this allocative inefficiency is large in New York City's rent control plan. It is therefore reasonable to conjecture that the unwinding of allocative distortions significantly contributed to Cambridge's residential price appreciation. Additional empirical analysis with rich micro-level attributes of residents, however, will be needed to shed further light on rent control's allocative consequences.

A key issue in the evaluation of price controls is the tradeoff between the surplus transferred from landlords to renters and the deadweight loss from quality or quantity undersupply. Viewed in this light, some portion of the price gains we measure at decontrolled properties are transfers from renters back to landlords. However, our analysis highlights the importance of another welfare consequence of price controls: the indirect effect on the desirability of housing in rent control-intensive locations. Our results indicate that the efficiency cost of Cambridge's rent control policy was large relative to the size of the transfer to renters. In particular, only 16 percent of the capitalized value of rent decontrol reflects a direct impact on decontrolled units, with the remainder due to indirect effects of rent control exposure on the amenity value of Cambridge residential units.

These findings are germane to economic analysis of housing market regulations and, more broadly, the impacts of other place-based policies. The mechanisms by which rent decontrol impacts never-controlled housing—increased maintenance, upgrading of local amenities, and potentially more efficient sorting of consumers to housing—are likely present in other settings involving residential housing. Our results provide evidence that residential spillovers are large and important in housing markets, and suggest that public policies related to housing should consider not only direct impacts but also indirect impacts on neighboring properties and residents.

Theory Appendix

We ground our empirical analysis in a stylized equilibrium model of the housing market that considers the relationship between rent control, neighborhood amenities, and house prices.

Neighborhoods. A city consists of $n = 1, ..., N$ neighborhoods. There is a continuum of locations in each neighborhood indexed by $\ell \in [0, 1]$. The pair (ℓ, n) refers to location ℓ in neighborhood n.

Landlords. Each location is owned by an absentee landlord who decides on the level of maintenance m. Maintenance includes inputs such as painting, upgrading, and repairs. These produce housing services according to the following increasing and concave technology

$$h = f(m).$$

While the model is static, we interpret housing services as a per-period flow variable. The price of housing services, p, is a per-period price.

The cost of maintenance is given by an increasing and convex function $c(m)$. The problem of the landlord is to choose a maintenance level m to maximize profits:

$$\max_{m} ph - c(m).$$

The first-order condition for an interior solution implies that maintenance is an increasing function of the price of housing services. Denote this function as $m^* = m(p)$ where $m'(p) > 0$.

Residents. Residents have preferences given by:

$$U(c, h) = Ac^{1-\alpha}h^{\alpha},$$

where c is a composite commodity and h is housing services. A is the total level of amenities in the neighborhood. The price of housing at location ℓ in neighborhood n is denoted $p_n(\ell)$, so a resident who lives at (ℓ, n) faces the budget constraint

$$c + p_n(\ell)h_n(\ell) = y,$$

where y denotes income. The only heterogeneity in the model comes from differences in income y

between residents. The outside utility for a resident with income y is denoted by $\bar{U}_y$.

Amenities depend on neighborhood attributes. To capture the most relevant dimensions for our study, we assume that amenities are increasing in the overall level of maintenance and income of residents as follows:

$$A_n = \int_0^1 (m_n(\ell) y_n(\ell))^\beta d\ell.$$

Here, $m_n(\ell)$ denotes the maintenance level at location ℓ in neighborhood n and $y_n(\ell)$ denotes the income of residents in neighborhood n residing at location ℓ and $\beta \in [0, \alpha)$. The equilibrium concept is based on spatial equilibrium, with free-entry and perfect mobility of residents.

Equilibrium definition. *An equilibrium is a triple $\langle y_n(\ell), p_n(\ell), h_n(\ell)\rangle$ where $y_n(\ell)$ is the resident income, $p_n(\ell)$ is the price, and $h_n(\ell)$ is the level of housing services for each neighborhood n and location ℓ such that*

- *each resident obtains at least their outside option,*
- *no resident wishes to move to another neighborhood or location within a neighborhood, and*
- *landlords maximize profits.*

Benchmark model. We impose particular functional forms to keep the model tractable. For the supply side, assume that housing services are produced by the linear technology $f(m) = m$ and the costs of maintenance are quadratic $c(m) = \frac{1}{2}m^2$. These assumptions imply that the optimal level of maintenance at each location is exactly equal to the price of housing services:

$$m^* = p.$$

The demand for housing services decreases with price:

$$h = \alpha y / p.$$

Next, we assume that the distribution of income among potential residents consists of N distinct levels of y, which we order from highest to lowest, $y_1 > \ldots > y_N$.

We first solve for the equilibrium without rent control as a baseline. We then consider the controlled equilibrium and develop implications for how prices, maintenance, and resident allocation will be affected by decontrol.

Equilibrium without rent control

We consider a symmetric equilibrium where all residents with income y_n live in neighborhood n. The log indirect utility V for a resident of neighborhood n at location ℓ is:

$$\ln V(p_n(\ell), y_n) = \ln(A_n) + \ln y_n - \alpha \ln p_n(\ell) + \ln((1-\alpha)^{1-\alpha}\alpha^{\alpha}).$$

Free-entry and perfect mobility of residents implies that in all locations ℓ in neighborhood n, each resident's utility is equal to $\bar{U}_{y_n}$. Hence, the price of housing services at each location ℓ is:

$$\ln(p_n(\ell)) = \frac{1}{\alpha}(\ln A_n + \ln y_n - \ln \bar{U}_{y_n} + \ln((1-\alpha)^{1-\alpha}\alpha^{\alpha}). \tag{2.5}$$

The value of neighborhood amenities comes from the fact that landlords optimally set the level of maintenance to the price of housing services and in the candidate equilibrium all residents of neighborhood n have income y_n. Therefore,

$$\ln(p_n(\ell)) = \frac{1}{\alpha}\left(\beta \int_0^1 \ln m_n(\ell)d\ell + (1+\beta)\ln y_n - \ln \bar{U}_{y_n} + \ln((1-\alpha)^{1-\alpha}\alpha^{\alpha})\right) = \ln(m_n(\ell)).$$

Symmetry among landlords implies that maintenance levels within a neighborhood are the same at each location location, so that

$$\ln(m_n) = \frac{1}{\alpha}\left(\beta \ln m_n + (1+\beta)\ln y_n - \ln \bar{U}_{y_n} + \ln((1-\alpha)^{1-\alpha}\alpha^{\alpha})\right).$$

This relationship captures the feedback between overall maintenance in the neighborhood and location specific maintenance choices. The maintenance levels in the uncontrolled economy m_n^u are

$$\ln(m_n) = \frac{1}{\alpha-\beta}((1+\beta)\ln y_n - \ln \bar{U}_{y_n} + \ln((1-\alpha)^{1-\alpha}\alpha^{\alpha}) \equiv \ln(m_n^u)$$

and prices are identical at all locations ℓ within neighborhood n. Using the expression for the level of maintenance, the price of housing p_n^u in neighborhood n in the economy without rent control is:

$$\ln(p_n) = \frac{1}{\alpha-\beta}((1+\beta)\ln y_n - \ln \bar{U}_{y_n} + \ln((1-\alpha)^{1-\alpha}\alpha^{\alpha}) \equiv \ln(p_n^u).$$

The pricing equation illustrates intuitive patterns under our parameter assumptions ($1 > \alpha > \beta \geq 0$). Prices are higher in neighborhoods when residents have more income and they are lower when residents have better outside options. Landlords invest more in response to exogenous improvements in neighborhood quality because more investment in the neighborhood raises amenities, which raises prices, and landlords set maintenance in response to prices.

Equilibrium with rent control

Let $RC_n \subset [0,1]$ denote the set of rent controlled locations in neighborhood n. Suppose that a fraction λ_n of locations are rent controlled and $1-\lambda_n$ are not. We first examine the pricing and maintenance decisions at controlled locations.

Rent controlled locations. Suppose the rent control authority sets prices at controlled locations $\bar{p}_n(\ell)$, and we assume that for each controlled location, the controlled price is less than the corresponding price in the uncontrolled economy, $\bar{p}_n(\ell) < p_n^u$.

This price will determine the level of maintenance according to the producer's first-order condition, which yields:

$$\bar{m}_n(\ell) = \bar{p}_n(\ell).$$

In turn, the amount of housing services at location ℓ is given by:

$$h_n(\ell) = f(\bar{m}_n(\ell)) = \bar{p}_n(\ell).$$

Uncontrolled locations. Spatial arbitrage determines the prices of uncontrolled locations and, hence, the arbitrage relation in equation (2.5) determines prices. Since $\bar{m}_n(\ell) = \bar{p}_n(\ell)$ and landlords are symmetric at uncontrolled locations, the level of amenities in the controlled economy is:

$$\ln A_n^c = \beta \left(\int_{\ell \in RC_n} \ln(\bar{p}_n(\ell)) d\ell + (1-\lambda_n) \ln m_n + \int_0^1 \ln y_n(\ell) d\ell \right).$$

As with the uncontrolled economy, we focus on the equilibrium where $y_n(\ell) = y_n$ for all uncontrolled locations ℓ. This yields:

$$\ln A_n^c = \beta \left(\int_{\ell \in RC_n} \left(\ln(\bar{p}_n(\ell)) + \ln y_n(\ell) \right) d\ell + (1-\lambda_n)(\ln m_n + \ln y_n) \right).$$

Since it is set by the rent control authority, the price of all controlled locations in neighborhood n may differ at each location, so we cannot further simplify the first term. For controlled locations, the income of a resident $y_n(\ell)$ depends on the way that residents are assigned to controlled housing. Let

$$\lambda_n \kappa_n^1 \equiv \int_{\ell \in RC_n} \ln(\bar{p}_n(\ell)) d\ell \quad \text{and} \quad \lambda_n k_n^2 \equiv \int_{\ell \in RC_n} \ln(y_n(\ell)) d\ell.$$

Since $\bar{p}_n(\ell) < p_n^u$, it is clear that

$$\kappa_n^1 < \ln(p_n^u).$$

While we do not explicitly model how residents are assigned to controlled housing, we assume that

$$\kappa_n^2 \leq \ln y_n,$$

which implies that the rationing mechanism imposed by rent control yields misallocation relative to the equilibrium in the uncontrolled economy.[48] The basis for this assumption is the following. If, prior to the implementation of rent control, the allocation were as in the symmetric equilibrium without rent control above, then once rent control is implemented, maintenance levels and hence housing services fall at controlled units (since landlords choose maintenance levels facing a regulated price). The combination of reduced rents and lower maintenance has one of two effects on incumbent residents: either they are sufficiently compensated by reduced rents so that they remain at their current locations, although the bundle of maintenance and amenities is not optimized for their income; or alternatively, they choose to relocate to areas with higher amenities and higher rents. In the latter case, they will be replaced by residents who prefer lower housing services, i.e., those with lower income levels. The average income at controlled locations will therefore weakly decline following the imposition of rent control.

As a result, amenities in neighborhood n in the presence of rent control are given by:

$$\ln A_n^c = \beta[\lambda_n(\kappa_n^1 + \kappa_n^2) + (1 - \lambda_n)(\ln m_n + \ln y_n)].$$

To compute the level of maintenance in uncontrolled locations in the presence of rent control, we follow similar steps to find:

$$\ln(m_n^c) \equiv \frac{1}{\alpha - \beta(1 - \lambda_n)} \left[\beta\lambda_n(\kappa_n^1 + \kappa_n^2) + (1 + \beta(1 - \lambda_n)) \ln y_n - \ln \bar{U}_{y_n} + \ln((1 - \alpha)^{1-\alpha}\alpha^{\alpha})\right].$$

[48]See, e.g., Suen (1989) for a canonical model of rationing in the presence of price controls. Bulow and Klemperer (2012) further investigate how consumer surplus is impacted by rationing and develop a model of rationing with rent-seeking.

We can write this in terms of the level of maintenance at uncontrolled locations in the economy without rent control:

$$\ln(m_n^c) = \frac{\alpha - \beta}{\alpha - \beta(1-\lambda_n)} \ln(m_n^u) + \frac{\beta}{\alpha - \beta(1-\lambda_n)} \Big(\lambda_n [\kappa_n^1 + \kappa_n^2] - \lambda_n \ln y_n \Big).$$

Summing up, since neighborhood amenities are a function of the maintenance of all units in a neighborhood and the income of residents, the supply of amenities at non-controlled locations in neighborhoods with rent controls—as well as maintenance and rents—is also impaired by rent control. This causes lower income residents to move into non-controlled locations. Hence, imposition of rent control causes inefficiently low maintenance and misallocation of residents at both controlled and non-controlled locations within a neighborhood.

The effect of rent control removal on rents, maintenance and resident allocation

Consider finally the impact of rent decontrol on prices at uncontrolled locations. To form this comparative static, we compare price levels in the economy without and with rent control:

$$\begin{aligned} \Delta \ln(p_n(\ell)) = \ln(p_n^u(\ell)) - \ln(p_n^c(\ell)) &= \frac{1}{\alpha}\Delta A_n = \underbrace{\frac{1}{\alpha}[A_n^u - A_n^c]}_{\text{indirect effect}} \\ &= \frac{\lambda_n \beta}{\alpha - \beta(1-\lambda_n)} \Bigg[\underbrace{[\ln(m_n^u) - \kappa_n^1]}_{\text{maintenance effect}>0} + \underbrace{[\ln y_n - \kappa_n^2]}_{\text{allocative effect}>0} \Bigg]. \end{aligned} \tag{2.6}$$

This expression shows that the end of rent control generates price impacts on uncontrolled locations through two channels in the model. For a given neighborhood, under rent control maintenance is inefficiently low and there are allocative inefficiencies due to the assignment of residents at controlled locations. This expression also illustrates three natural comparative statics. When a neighborhood has a higher fraction of locations that are controlled (λ_n increases), the change in prices for locations without rent control increases. As κ_n^1 increases (as would be expected when the prices of controlled locations are further depressed from their market values), the change in the price of uncontrolled locations due to the elimination of rent control also increases. Moreover, when there is greater misallocation due to the rent control (κ_n^2 decreases), the elimination of rent control further increases prices.

The price impact due to the end of rent control for formerly controlled locations involves an additional term which can be decomposed as follows:

$$\ln(p_n^u(\ell)) - \ln(\bar{p}_n^c(\ell)) = \left(\underbrace{\ln(p_n^u(\ell)) - \ln(p_n^c(\ell))}_{\text{indirect effect}}\right) + \left(\underbrace{\ln(p_n^c(\ell)) - \ln(\bar{p}_n^c(\ell))}_{\text{direct effect}}\right). \tag{2.7}$$

The first term, the *indirect effect*, is the price change for uncontrolled locations due to the end of rent control, which is in turn due to maintenance and allocative effects as in equation (2.6). The second term, the *direct effect*, is the price change in a controlled economy going from a rent controlled location to an uncontrolled location. For a formerly controlled location. the direct effect of the end of rent control is larger when the controlled price at the location is further depressed. The following proposition summarizes the relevant considerations from this model:

Proposition 1. *When rent control ends, the price change for uncontrolled locations is greater for neighborhoods*

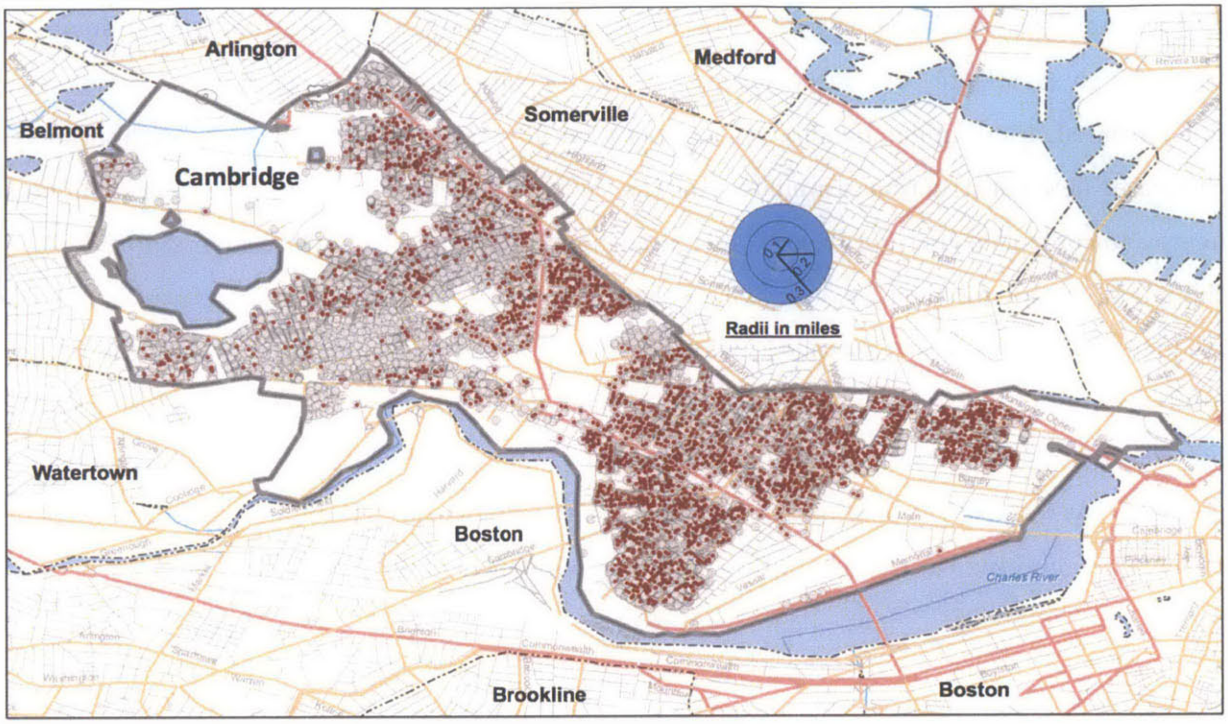
- *with a larger fraction of locations with rent control* ($\lambda \uparrow$),
- *where the price of controlled locations is further depressed from their market price* ($\kappa^1 \downarrow$),
- *where there is greater misallocation of resident types relative to the types in the uncontrolled economy* ($\kappa^2 \downarrow$).

Furthermore, when rent control ends, controlled locations experience an additional price increase due to the direct effect of decontrol.

This model shows the difficulty involved in distinguishing between direct and indirect effects at decontrolled locations. When rent control ends, there is a direct price effect due to the formerly controlled location being priced by the market. However, there is also an indirect effect as neighborhood amenities improve due to increases in maintenance and the income of residents, leading to higher prices. This in turn leads a landlord to invest in additional maintenance. For empirical purposes, at decontrolled locations, the direct and indirect channels cannot be readily distinguished because each affects the equilibrium level of the other.

The model's simplicity also imposes some limitations for our setting. First, the price of housing services is an abstraction that allows for no distinction between house prices and rents, which might be especially relevant in a dynamic setting. The model does not therefore allow for realistic dynamics to capture expectations of neighborhood appreciation and the option value of ownership. Second, amenities within a neighborhood are assumed to be pure public goods, so residents have no desire to substitute between locations within a neighborhood. If housing services were instead differentiated, there might be substitution between different locations within a neighborhood. In this case, new construction stimulated by the end of rent control might have a price impact at nearby uncontrolled housing (due to increased housing supply). Third, residents are identical in the model except for their income: within a neighborhood, all residents at uncontrolled locations (though not generally at controlled locations) have the same level of income and, due to spatial arbitrage, obtain the same utility. Finally, the model focuses on one housing market, and does not consider neighboring markets that do not have rent control. Although it is not modeled, it is possible that residents at previously controlled locations move out of Cambridge, and that residents in these neighboring towns move into Cambridge with the end of rent control.

Figure 1. The Geography of Residential Properties and Rent Control in Cambridge, Massachusetts



Notes. Cambridge residential properties as of 2008 are marked with gray circles. Map-lots that were rent-controlled as of 1994 are overlaid with red circles. Concentric circles in the top right depict radii of 0.1, 0.2, and 0.3 miles.

Figure 2. Residential Turnover in Cambridge Controlled relative to Never-Controlled Units, 1992 - 2000

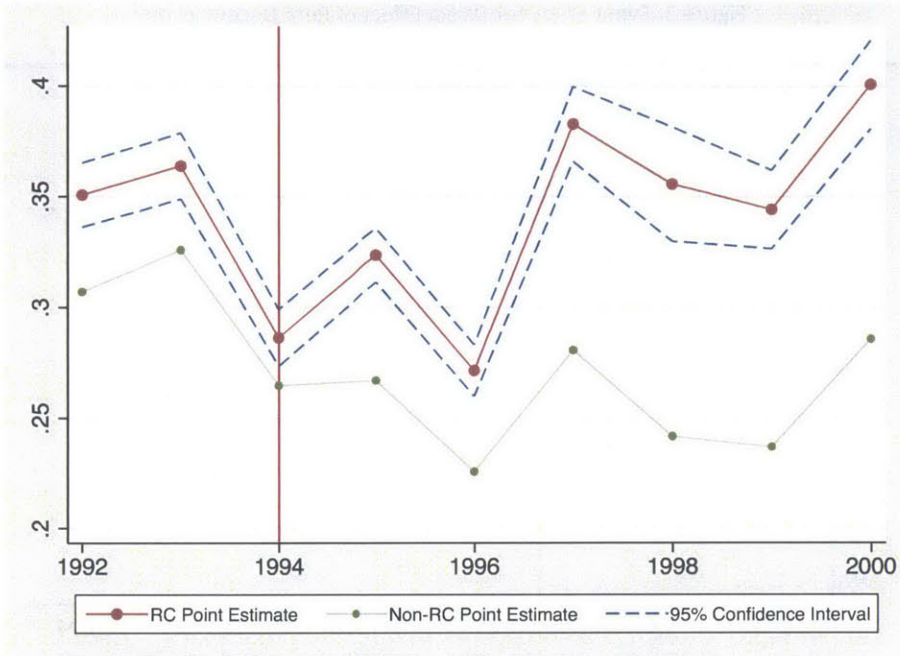
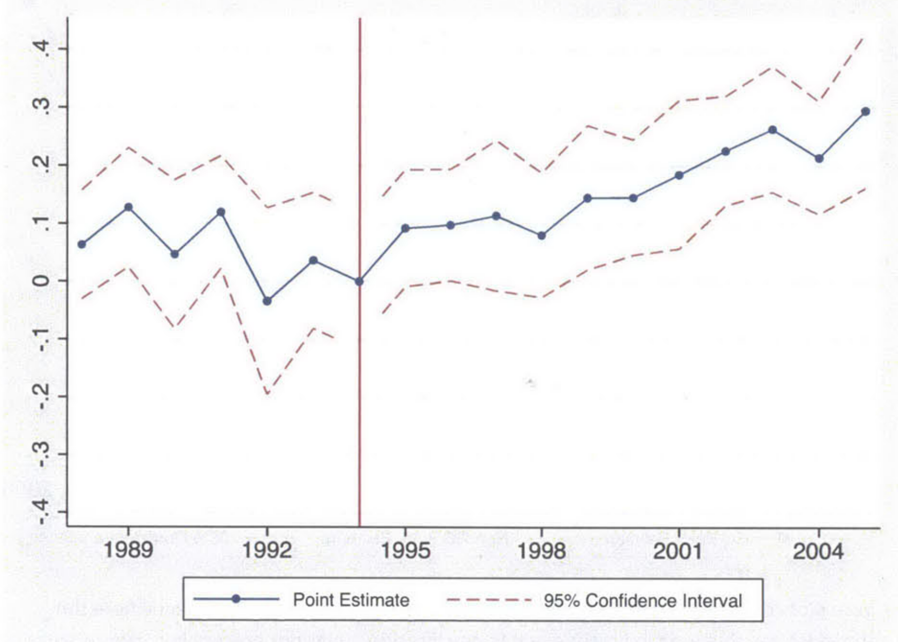


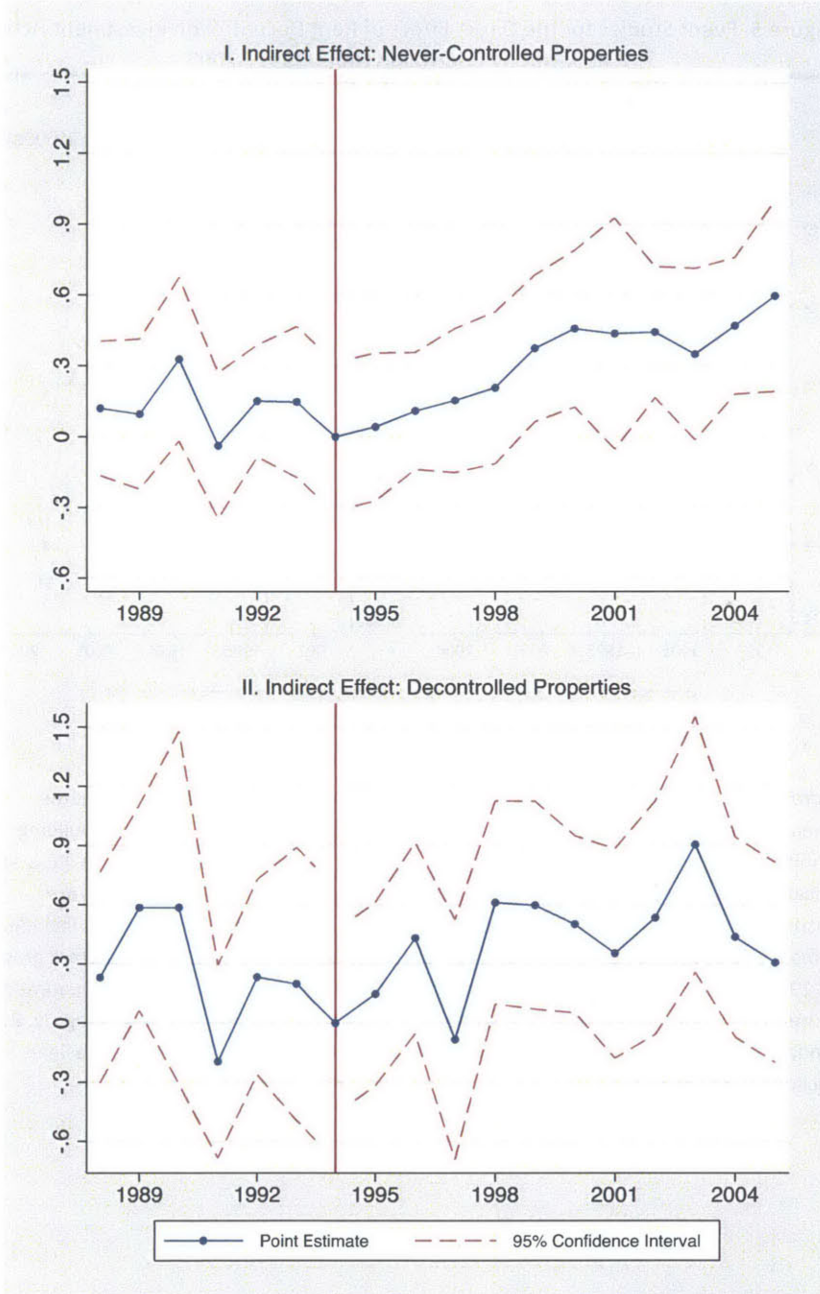
Figure plots coefficients on RC x Year variables from an event-study regression where the dependent variable is an indicator equal to one if resident was not present in current Cambridge unit in prior year (and zero otherwise). RC is an indicator for a location that was rent controlled in 1994. This specification includes an RC main effect, year controls, structure type dummies, and geographic fixed effects for the 91 block groups in the 1990 Census containing addresses listed in the Cambridge City Census. 95% confidence intervals are constructed from robust standard errors clustered by block group. The vertical line in 1994 indicates the year preceding rent control removal.

Figure 3. Event Study for Direct Effect of Rent Decontrol on Transaction Prices of Decontrolled Units, 1988 - 2005



Figures plot RC x Year coefficients from event study regressions where dependent variable is log sale price, winsorized to the first percentile separately for houses and condominiums. RC is an indicator for a location that was rent controlled in 1994. Regression also includes year dummies, block group fixed effects, structure type main effects and quadratic time trends, and controls for property characteristics: total rooms, bathrooms, bedrooms, interior square feet, lot size and its square, a dummy for lot size zero, log property age and its square, and a dummy for property age missing, all interacted with structure type dummies. Robust standard errors are clustered by block group. Vertical line in 1994 designates the year preceding rent decontrol.

Figure 4. Event Study for Indirect Effect of Rent Decontrol on Transaction Prices of Never-Controlled and Decontrolled Properties, 1988 - 2005



Figures plot RCI x Year coefficients from event study regressions of log sale prices, winsorized to the first percentile for houses and condominiums. RCI is calculated over a 0.20 mile radius. Panels A and B are estimated using never-controlled and formerly controlled properties. See Figure 3 notes for specification details. Robust standard errors are clustered by block group. Vertical line in 1994 designates the year preceding rent decontrol.

Figure 5. Event Studies for the Direct Effect of Rent Decontrol on Investment Activity at Formerly Controlled Units, 1991 - 2005

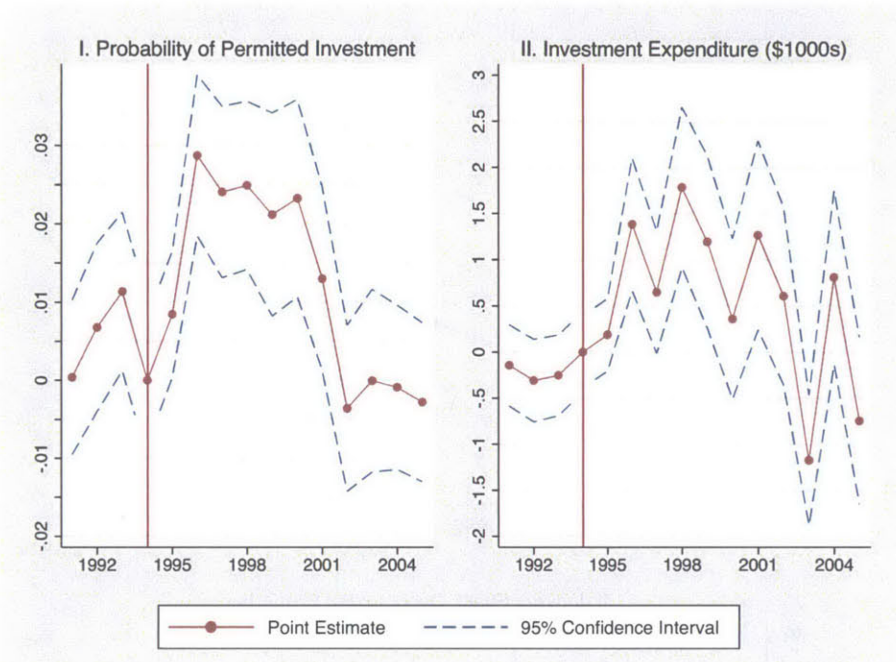


Figure plots coefficients on RC x Year variables from event study regressions where the dependent variable is (left panel) an indicator for whether a structure received a building permit and (right panel) the permitted expenditure at a structure (including zeros). RC is an indicator for a location that was rent controlled in 1994. Investment expenditures are winsorized by structure type and year to the 99.5th percentile. Both specifications include an RC main effect, year fixed effects, geographic fixed effects for the 89 Cambridge block groups in the 1990 Census containing assessed properties, structure type indicators, and a quadratic in the number of units in condominium structures. 1994 is the omitted RC x Year category. Robust standard errors are clustered by block group. The vertical line in 1994 indicates the year preceding rent decontrol.

Table 1. Turnover at Cambridge Residential Locations, 1992-2000

Dependent Variable: Indicator Equal to One if Resident was not at Location in Prior Year

	All Properties (1)	Houses (2)	Condominiums (3)	Apartments (4)
Mean of Dependent	0.269	0.232	0.297	0.335
	(0.197)	(0.178)	(0.209)	(0.223)
RC	-0.003	0.073***	-0.035**	-0.056**
	(0.008)	(0.008)	(0.016)	(0.026)
RC x Post	0.054***	0.025***	0.076***	0.057**
	(0.008)	(0.008)	(0.022)	(0.025)
N	310,949	172,996	70,558	67,395

Dependent variable is an indicator equal to one if resident was not present in current unit in prior year (and zero otherwise). RC is an indicator for a location that was rent controlled in 1994 and Post is an indicator for year 1995 and after. All specifications include year controls, structure type dummies, and geographic fixed effects for the 91 block groups in the 1990 Census containing addresses listed in the Cambridge City Census. Robust standard errors clustered by block group are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B2. Effects of Rent Decontrol and Rent Control Intensity on Assessed Values by Structure Type Dropping Block Groups Bordering non-Cambridge Properties

	(1)	(2)	(3)	(4)	(5)	(6)
			I. Houses			
RC x Post	0.091***	0.073***	0.023	-0.007	0.080***	0.034
	(0.016)	(0.013)	(0.042)	(0.024)	(0.016)	(0.044)
RCI x Post	0.420***	0.317***			0.434***	
	(0.089)	(0.095)			(0.085)	
Non-RC x RCI x Post			0.403***	0.297***		0.423***
			(0.088)	(0.095)		(0.084)
RC x RCI x Post			0.597***	0.525***		0.556***
			(0.154)	(0.121)		(0.159)
H_0: No Spillovers			0.000	0.000		0.000
H_0: Spillovers Equal			0.123	0.003		0.310
R-squared	0.847	0.850	0.847	0.850	0.850	0.850
N	12,407	12,407	12,407	12,407	11,415	11,415
			II. Condominiums			
RC x Post	0.335***	0.343***	0.397**	0.269**	0.322***	0.218
	(0.050)	(0.042)	(0.166)	(0.134)	(0.046)	(0.166)
RCI x Post	0.230	0.691**			0.101	
	(0.171)	(0.283)			(0.197)	
Non-RC x RCI x Post			0.283	0.571**		-0.040
			(0.185)	(0.276)		(0.126)
RC x RCI x Post			0.123	0.744**		0.213
			(0.321)	(0.326)		(0.358)
H_0: No Spillovers			0.290	0.042		0.819
H_0: Spillovers Equal			0.670	0.594		0.530
R-squared	0.701	0.710	0.701	0.710	0.712	0.712
N	14,140	14,140	14,140	14,140	10,131	10,131
Block group FEs	y	y	y	y	y	y
Tract Trends	-	y	-	y	-	-
Excluding Converted	-	-	-	-	y	y

Notes. Dependent variable is log assessed value. Assessed values are from 1994 and 2004. RCI is calculated based on a radius of 0.20 miles. RCI is demeaned so that the main rent control effect is the price differential for the average rent controlled property. Post is an indicator for year equal to 2004. Year fixed effects and block group fixed effects are included in all regressions. Block group fixed effects correspond to each of the 88 Cambridge block groups containing assessed properties using 1990 Census boundaries. Tract trends are tract x post dummies for each of 30 tracts from the 1990 Census. Two-family and three-family house dummies are included in the top panel. Test for No Spillovers report p-vaues from tests that RCI x Post or Non-RC x RCI x Post and RC x RCI x Post coefficients are zero. Test for Spillovers Equal reports p-value for hypothesis that these latter two coefficients are equal. Robust standard errors clustered by 1990 block group are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3. Effects of Rent Decontrol on Assessed Values
Dependent Variable: Log of Assessed Property Value (1994, 2004)

	(1)	(2)	(3)	(4)
RC	-0.504***	-0.504***	-0.515***	
	(0.075)	(0.052)	(0.052)	
RC x Post	0.217***	0.227***	0.249***	0.221***
	(0.039)	(0.037)	(0.034)	(0.040)
Block Group FEs	No	Yes	Yes	No
Tract Trends	No	No	Yes	Yes
Map-lot FEs	No	No	No	Yes
R-squared	0.605	0.759	0.763	0.938

N = 32,980. Sample is all assessed Cambridge houses and condominium properties in 1994 and 2004. RC is an indicator for a location that was rent controlled in 1994 and Post is an indicator for year equal to 2004. Year fixed effects and structure-type dummies are included in all regressions. Block group fixed effects correspond to the 89 Cambridge block groups in the 1990 Census containing assessed properties. Tract trends are tract x Post dummies for each of 30 tracts from the 1990 Census. Map-lot FEs are dummy variables for each of the 9,497 residential parcels in Cambridge. Map-lot FEs absorb the RC main effect in column 4. Robust standard errors clustered by 1990 block group are in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 4. Effects of Rent Decontrol and Rent Control Intensity on Assessed Values
Dependent Variable: Log of Assessed Property Value (1994, 2004)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
RC	-0.440***	-0.484***	-0.503***		-0.232	-0.217	
	(0.057)	(0.050)	(0.052)		(0.188)	(0.184)	
RC x Post	0.175***	0.196***	0.233***	0.208***	0.202*	0.174	0.132
	(0.038)	(0.036)	(0.034)	(0.040)	(0.114)	(0.107)	(0.114)
RCI	-0.581*	-0.792	-0.938*				
	(0.325)	(0.479)	(0.494)				
RCI x Post	0.328**	0.258*	0.545***	0.475***			
	(0.136)	(0.138)	(0.191)	(0.180)			
Non-RC x RCI					-0.568	-0.686	
					(0.546)	(0.561)	
Non-RC x RCI x Post					0.281*	0.514**	0.415*
					(0.168)	(0.227)	(0.220)
RC x RCI					-1.211**	-1.416**	
					(0.535)	(0.555)	
RC x RCI x Post					0.249	0.651***	0.607**
					(0.215)	(0.231)	(0.256)
Block Group FEs	No	Yes	Yes	No	Yes	Yes	No
Tract Trends	No	No	Yes	Yes	No	Yes	Yes
Map-Lot FEs	No	No	No	Yes	No	No	Yes
H_0: RCI x Post coeffs equal					0.909	0.598	0.514
R-squared	0.611	0.761	0.765	0.938	0.764	0.767	0.938

Number of observations is 32,980. RCI is calculated over a 0.20 mile radius and de-meaned. RC is an indicator for a location that was rent controlled in 1994 and Post is an indicator for year equal to 2004. RC and RC x RCI main effects are absorbed by map lot fixed effects (FEs) in columns 4 and 7. Year fixed effects and structure type dummies are included in all regressions. Block group FEs correspond to the 89 Cambridge block groups in the 1990 Census containing assessed properties. Tract trends are tract x Post dummies for each of 30 tracts from the 1990 Census. Map-lot FEs are dummy variables for each of the 9,497 residential pracels in Cambridge. Test of the equality of the RCI x Post coefficients reports p-values from tests that Non-RC x RCI x Post and RC x RCI x Post coefficients are equal. Robust standard errors clustered by 1990 block group are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5. Effects of Rent Decontrol and Rent Control Intensity on Assessed Values by Structure Type

Dependent Variable: Log of Assessed Property Value (1994, 2004)

	(1)	(2)	(3)	(4)	(5)	(6)
			I. Houses			
RC x Post	0.065***	0.045***	0.024	0.035	0.035	0.035
	(0.011)	(0.016)	(0.023)	(0.036)	(0.023)	(0.032)
RCI x Post	0.205*	0.200				
	(0.103)	(0.144)				
Non-RC x RCI x Post			0.194*	0.197	0.192**	0.190
			(0.103)	(0.142)	(0.095)	(0.135)
RC x RCI x Post			0.315**	0.227	0.232*	0.231
			(0.130)	(0.196)	(0.128)	(0.181)
H_0: RCI x Post coeffs equal			0.080	0.782	0.553	0.675
R-squared	0.855	0.984	0.855	0.984	0.858	0.983
N	16,239	16,239	16,239	16,239	14,917	14917
			II. Condominiums			
RC x Post	0.354***	0.345***	0.361***	0.276**	0.235*	0.236*
	(0.038)	(0.037)	(0.135)	(0.131)	(0.132)	(0.136)
RCI x Post	0.669**	0.492**				
	(0.256)	(0.211)				
Non-RC x RCI x Post			0.678**	0.397	0.443**	0.454**
			(0.308)	(0.258)	(0.205)	(0.206)
RC x RCI x Post			0.648**	0.569**	0.722**	0.724**
			(0.291)	(0.266)	(0.323)	(0.328)
H_0: RCI x Post coeffs equal			0.925	0.586	0.398	0.429
R-squared	0.714	0.889	0.714	0.889	0.725	0.89
N	16,741	16,741	16,741	16,741	11,778	11778
Block Group FEs	Yes	No	Yes	No	Yes	No
Map-Lot FEs	No	Yes	No	Yes	No	Yes
Tract Trends	Yes	Yes	Yes	Yes	Yes	Yes
Excluding Converted Structures	No	No	No	No	Yes	Yes

RCI is calculated over a 0.20 mile radius. RC is an indicator for a location that was rent controlled in 1994 and Post is an indicator for year equal to 2004. In specifications that include RC, RCI, Non-RC x RCI or RC x RCI interacted with Post, main effects of these variables are included but not tabulated. Year fixed effects (FEs) and structure type dummies are included in all regressions. Block group FEs correspond to the 89 Cambridge block groups in the 1990 Census containing assessed properties. Map-Lot FEs are a set of dummies for each residential parcel (8,453 for houses, 1,450 for condominiums). Tract trends are tract x Post dummies for each of 30 tracts from the 1990 Census. Final columns exclude units that change usage categories between 1994 and 2004. Robust standard errors clustered by 1990 block group are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6. Effect of Rent Decontrol and Rent Control Intensity on Assessed Values for Various Rent Control Intensity Measures

Dependent Variable: Log of Assessed Property Value (1994, 2004)

	I. RCI Defined Over Varying Geographies			
	(1)	(2)	(3)	(4)
	0.10 miles	0.20 miles	0.30 miles	Census Block Group
	I. Varying the Geographies Used to Measure RCI			
RC x Post	0.132	0.132	0.149	0.128
	(0.089)	(0.114)	(0.125)	(0.098)
Non-RC x RCI x Post	0.185	0.415*	0.477*	0.095
	(0.143)	(0.220)	(0.245)	(0.177)
RC x RCI x Post	0.377**	0.607**	0.646**	0.318
	(0.183)	(0.256)	(0.281)	(0.228)
H_0: RCI x Post coeffs equal	0.379	0.514	0.594	0.367
Std Dev of RCI measure	0.192	0.165	0.145	0.179
Geographic FEs	Map-Lot	Map-Lot	Map-Lot	Map-Lot
	II. Varying the Weighting Scheme Used to Measure RCI			
	(1)	(2)	(3)	(4)
	λ = -12	λ = -9	λ = -6	λ = -3
Weight @ 0.1 mile/ 0.01 mile	0.34	0.44	0.58	0.76
Weight @ 0.2 mile/ 0.01 mile	0.10	0.18	0.32	0.57
Weight @ 0.3 mile/ 0.01 mile	0.03	0.07	0.18	0.42
RC x Post	0.124	0.129	0.136	0.134
	(0.124)	(0.136)	(0.155)	(0.208)
Non-RC x RCI x Post	0.407**	0.499**	0.598**	0.629**
	(0.193)	(0.223)	(0.261)	(0.264)
RC x RCI x Post	0.580**	0.665**	0.758**	0.802*
	(0.243)	(0.280)	(0.345)	(0.457)
H_0: RCI x Post coeffs equal	0.504	0.555	0.617	0.683
Std Dev of RCI measure	0.190	0.179	0.160	0.117
Geographic FEs	Map-Lot	Map-Lot	Map-Lot	Map-Lot
Tract x Yr FEs	Yes	Yes	Yes	Yes

N=32,980. In panel I, RCI is calculated over geographies reported in column headings. In panel II, RCI is calculated using an exponential decay weighting scheme (see equation (3) in the text). RC is an indicator for a location that was rent controlled in 1994 and Post is an indicator for year equal to 2004. Year fixed effects (FEs), tract x year dummies, and individual Map-Lot FEs are included in all specifications. Robust standard errors clustered by 1990 block group (89 groups) are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7. Effects of Rent Decontrol and Rent Control Intensity on Transaction Prices, 1988 - 2005
Dependent Variable: Log Sale Price

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
RC	-0.305***	-0.204***	-0.193***	-0.189***	-0.185***	-0.166***	-0.161***
	(0.043)	(0.024)	(0.024)	(0.025)	(0.024)	(0.025)	(0.024)
RC x Post	0.060*	0.106***	0.086***	0.087***	0.079***	0.079***	0.068***
	(0.030)	(0.026)	(0.027)	(0.026)	(0.025)	(0.025)	(0.024)
RCI				-0.510*	-0.494		
				(0.305)	(0.317)		
RCI x Post				0.205***	0.166*		
				(0.056)	(0.098)		
Non-RC x RCI						-0.305	-0.276
						(0.274)	(0.275)
Non-RC x RCI x Post						0.197***	0.132
						(0.067)	(0.089)
RC x RCI						-0.884**	-0.883**
						(0.360)	(0.368)
RC x RCI x Post						0.246*	0.246
						(0.146)	(0.177)
Block Group FEs	No	Yes	Yes	Yes	Yes	Yes	Yes
Property Characteristics	No	Yes	Yes	Yes	Yes	Yes	Yes
Quadratic Tract Trends	No	No	Yes	No	Yes	No	Yes
H_0: RCI x Post coeffs equal	No	No	No	No	No	0.773	0.512
R-squared	0.318	0.674	0.681	0.675	0.682	0.678	0.684

N=14,789 Cambridge house and condominium properties transacted during 1988 through 2005. Prices are winsorized by structure type at the first percentile. RCI is defined over a 0.20 mile radius and demeaned. RC is an indicator for a location that was rent controlled in 1994 and Post is an indicator for year is 1995 or afterwards. All specifications include year of sale dummies and structure type dummies. Property Characteristics, each interacted with structure type, include number of total rooms, bathrooms, bedrooms, interior square footage, a dummy variable for zero lot size, a quadratic in lot size, a dummy variable for missing year built, a quadratic in the log age of the structure, and a quadratic time trend for each structure type. Block group fixed effects correspond to each of the 88 Cambridge block groups in the 1990 Census containing transacted properties. Columns 5 and 7 include quadratic tract trends for each of 30 Census tracts. Test for No Spillovers reports p-values from tests that RCI x Post or Non-RC x RCI x Post and RC x RCI x Post coefficients are zero. Test for Spillovers Equal reports p-values from tests that these latter two coefficients are equal. Robust standard errors clustered by 1990 block group are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 8. Estimated Direct and Indirect Contributions of Rent Decontrol to Changes in Cambridge Assessed Residential Property Values, 1994 to 2004 (in millions of 2008 dollars)

	I. Assessed Housing Values				II. Estimated Effects of Decontrol on Housing			
					Increase in Value ($)		Increase in Value (%)	
	1994 Assessed (mil$)	2004 Assessed (mil$)	Change 1994 - 2004 Δ $	Δ %	Direct Effect	Indirect Effect	Direct Effect	Indirect Effect
				a. Decontrolled Units				
Houses	$267	$760	$493	185%	$94	$149	18%	29%
Condominiums	$518	$1,746	$1,228	237%	$216	$390	19%	34%
All	$785	$2,507	$1,722	219%	$310	$539	19%	33%
				b. Never-Controlled Units				
Houses	$2,961	$7,320	$4,359	147%	n/a	$822	n/a	13%
Condominiums	$1,017	$2,699	$1,683	166%	n/a	$306	n/a	13%
All	$3,978	$10,020	$6,042	152%	n/a	$1,128	n/a	13%
				c. All Units				
Houses	$3,229	$8,081	$4,852	150%	$94	$971	1%	14%
Condominiums	$1,535	$4,446	$2,911	190%	$216	$696	6%	20%
All	$4,763	$12,526	$7,763	163%	$310	$1,667	3%	16%

Assessed values are from the 1995 and 2005 Cambridge Assessor's databases, reflecting property valuations as of 1994 and 2004, respectively. Counterfactual log property values are estimated separately for houses and condos using the specification in column 7 of Table 4. Counterfactuals for RCI effects subtract Non-RC x RCI x Post and RC x RCI x Post effects and counterfactuals for the direct effect of decontrol subtract RC x Post effects from actual log property values. Aggregate effects in 2008 dollars are calculated by summing exponentiated counterfactual log property values.

Table A1. Cambridge Residential Building Permitting Activity, 1991 through 2004: Permits Issued and Permitted Expenditures

	Houses				Condominiums			
	Never Controlled		Decontrolled		Never Controlled		Decontrolled	
	1991-1994	1995-2004	1991-1994	1995-2004	1991-1994	1995-2004	1991-1994	1995-2004
	I. Permits Issued							
Permits	1,507	4,385	259	694	247	852	185	672
Pr(Permit)	0.030	0.035	0.029	0.031	0.014	0.019	0.011	0.016
E[Units | Permit]	1.72	1.72	2.54	2.81	12.06	10.95	15.69	16.34
	II. Annual Expenditure (1,000s of 2008 dollars)							
Total	14,044	29,954	1,588	3,486	3,723	7,595	1,451	4,435
Per Unit	1.11	2.37	0.72	1.57	0.82	1.67	0.34	1.05
	III. Annual Expenditure per Permitted Unit (1,000s of 2008 dollars)							
Mean	37.3	68.3	24.5	50.2	60.3	89.1	31.4	66.0
SD	164.5	178.0	46.8	105.6	190.2	338.4	118.1	269.6
Median	10.3	18.0	8.3	13.8	12.4	19.3	11.2	19.2
Min	0.1	0.1	0.4	0.3	0.5	0.3	0.4	0.4
Max	5,675.5	4,365.5	451.2	1,208.9	2,121.2	6,589.3	1,480.1	4,450.3

Data source is the universe of Cambridge Inspectional Services permits issued during 1991 through 2004. If a structure receives multiple permits in a given year, we sum these expedentitures and treat them as a single permit. When calculating units permitted or expenditures per unit in a year, we attribute the structure's permitted status and expenditures to only one unit. Expenditures are converted to real 2008 dollars using the Consumer Price Index for All Items Less Shelter for All Urban Consumers, Series Id: CUUR0000SA0L2, Not Seasonally Adjusted.

Table A2. Property Conversions, 1994-2004:
Status in 1994 of Units that Were Designated as Houses and Condominiums in 2004

	2004 Houses			2004 Condominiums		
1994 Structure Type	All Houses	Formerly Controlled	Never Controlled	All Condo-miniums	Formerly Controlled	Never Controlled
Same as 2004	13,480 (97.3%)	1,567 (89.9%)	11,913 (98.3%)	7,085 (74.1%)	3,507 (76.2%)	3,578 (72.1%)
Converted from	381 (2.7%)	177 (10.1%)	204 (1.7%)	2,476 (25.9%)	1,093 (23.8%)	1,383 (27.9%)
Houses				1,058 (11.1%)	151 (3.3%)	907 (18.3%)
Condominiums	20 (0.1%)	3 (0.2%)	17 (0.1%)			
Apartments	153 (1.1%)	115 (6.5%)	38 (0.3%)	647 (6.8%)	599 (13%)	48 (1%)
Other Residential	50 (0.4%)	35 (2%)	15 (0.1%)	347 (3.6%)	284 (6.2%)	63 (1.3%)
Non-Residential	158 (1.1%)	24 (1.4%)	134 (1.1%)	424 (4.4%)	59 (1.3%)	365 (7.4%)
Total	13,861	1,744	12,117	9,561	4,600	4,961

Counts and conversion rates are calculated from Cambridge Assessor's databases, reflecting property characteristics as of 1994 and 2004. The "Other Residential" category includes structures zoned as boarding houses, mixed use, or multiple houses on a single parcel.

Table A3. Descriptive Statistics for Population and Residential Units and Structures for Various Geographies

	Mean	Std Dev	Min	Max	Median
			I. Census Blocks		
Area (sq miles)	0.01	0.02	0.00	0.53	0.00
1990 Census Population	135.05	162.71	0.00	2833.00	99.00
2001 Residential Units	62.77	58.71	0.00	441.00	45.00
1994 Rent Control Units	22.92	34.48	0.00	236.00	11.00
2001 Residential Structures	18.53	12.08	0.00	81.00	16.00
1994 Rent Control Structures	4.08	3.77	0.00	21.00	3.00
Count of Blocks			587		
			II. Census Block Groups		
Area (sq miles)	0.07	0.07	0.01	0.56	0.05
1990 Census Population	986.17	506.00	98.00	3093.00	836.00
2001 Residential Units	428.15	253.62	23.00	1418.00	387.00
1994 Rent Control Units	155.75	155.19	6.00	854.00	107.00
2001 Residential Structures	122.93	58.53	9.00	382.00	124.00
1994 Rent Control Structures	27.26	16.30	3.00	61.00	24.00
Count of Block Groups			89		
			III. Census Tracts		
Area (sq miles)	0.22	0.17	0.05	0.72	0.16
1990 Census Population	3144.73	1291.67	1736.00	7123.00	2650.00
2001 Residential Units	1291.68	510.60	336.00	2984.46	1244.07
1994 Rent Control Units	470.77	341.71	101.00	1534.00	379.50
2001 Residential Structures	365.00	149.06	117.00	860.00	338.50
1994 Rent Control Structures	80.90	30.41	27.00	156.00	73.00
Count of Tracts			30		
			IV. 0.2 mile radius		
Area (sq miles)	0.13	-	0.13	0.13	0.13
1990 Census Population	3160.48	1765.02	0.00	15796.90	2935.48
2001 Residential Units	1141.15	573.10	5.00	3427.54	1066.16
1994 Rent Control Units	422.34	330.59	0.00	1702.00	376.00
2001 Residential Structures	348.40	116.72	1.00	676.00	351.00
1994 Rent Control Structures	80.15	46.52	0.00	180.00	77.00
Count of Maplots			10,968		

Panels I through III provide statistics for all census geographics containing at least one assessed residential housing structure. Panel IV provides statistics for the universe of 0.2 mile radius geographies, centered at each residential housing structure.

Table A4. Descriptive Statistics for Transacted Properties

	Houses				Condominiums			
	Never Controlled		Decontrolled		Never Controlled		Decontrolled	
	1988-1994	1995-2005	1988-1994	1995-2005	1988-1994	1995-2005	1988-1994	1995-2005
log Price	12.84	13.26	12.59	13.03	12.56	12.81	12.20	12.57
	(0.69)	(0.74)	(0.67)	(0.67)	(0.51)	(0.55)	(0.56)	(0.55)
Total Rooms	9.16	9.40	10.24	10.27	4.77	5.03	4.40	4.41
	(3.33)	(3.43)	(3.57)	(3.67)	(1.53)	(1.91)	(1.60)	(1.55)
Bedrooms	4.05	4.10	4.56	4.61	2.00	2.12	1.68	1.75
	(1.69)	(1.72)	(1.80)	(1.85)	(0.78)	(0.96)	(0.70)	(0.81)
Bathrooms	2.77	2.81	2.93	2.91	1.57	1.63	1.17	1.24
	(0.94)	(0.95)	(0.87)	(0.85)	(0.67)	(0.75)	(0.44)	(0.52)
Interior sq. ft.	2363.41	2387.34	2408.88	2409.76	1202.67	1269.57	927.85	949.69
	(1131.25)	(1071.66)	(920.96)	(902.49)	(834.76)	(819.75)	(434.02)	(449.68)
Has Lot (y/n)	0.99	0.99	0.99	0.99	0.02	0.04	0.04	0.03
	(0.11)	(0.09)	(0.09)	(0.09)	(0.14)	(0.19)	(0.18)	(0.17)
Lot Size sq. ft.	4211.71	4253.09	3320.15	3462.02	113.24	157.66	191.18	151.38
	(3433.26)	(3437.64)	(1964.22)	(2031.41)	(1595.75)	(1145.06)	(1222.04)	(1148.19)
Year Built	1903.25	1903.31	1890.81	1892.71	1944.51	1935.16	1915.12	1916.42
	(36.93)	(37.81)	(24.67)	(24.94)	(44.72)	(45.58)	(27.94)	(30.86)
N	1,624	2,599	255	336	2,138	3,626	1,446	2,765

Sample includes Cambridge houses and condominiums transacted during 1988 through 2005. Sales price, winsorized by structure type to the 1st percentile, are converted to real 2008 dollars using the Consumer Price Index for All Items Less Shelter for All Urban Consumers, Series Id: CUUR0000SA0L2, Not Seasonally Adjusted. Has Lot indicates whether property has a non-zero lot size.

Table A5. Comparison of Estimated Relationship between Rent Control Status, Rent Control Intensity, and Transacted Prices vs. Assessed Values for Units Transacted in 1994 and 2004
Dependent Variable: Log of Transacted or Assessed Price

	Transacted Prices			Assessed Values: Transacted Units		
	(1)	(2)	(3)	(4)	(5)	(6)
			I. Houses			
RC x Post	0.199	0.127	0.352**	0.114	0.059	0.194
	(0.124)	(0.125)	(0.163)	(0.078)	(0.081)	(0.137)
RCI x Post		0.606**			0.452**	
		(0.294)			(0.189)	
Non-RC x RCI x Post			0.736***			0.522***
			(0.278)			(0.193)
RC x RCI x Post			-0.539			-0.172
			(0.828)			(0.615)
N	685	685	685	652	652	652
			II. Condominiums			
RC x Post	0.163**	0.085	0.073	0.168**	0.133*	0.122*
	(0.072)	(0.068)	(0.063)	(0.071)	(0.074)	(0.069)
RCI x Post		0.512**			0.255	
		(0.200)			(0.201)	
Non-RC x RCI x Post			0.406			0.110
			(0.280)			(0.294)
RC x RCI x Post			0.709**			0.516
			(0.291)			(0.366)
N	937	937	937	7,897	7,897	7,897

Samples includes houses (Panel I) and condominiums (Panel II) that transacted in 1994 and 2004. Regression models follow column 5 of Table 4. In columns 4-6, the dependent variable is the assessed value of any unit that is on a map lot at which at least one unit transacted in the given year. The number of observations for houses is larger in columns 1-3 than in columns 4-6 because a unit may transact more than once per year. The number of observations for condominimums in columns 4-6 is larger than in columns 1-3 because condominium structures contain multiple units. We observe the market price for transacted units and the assessed price for all units in the structure but cannot determine which specific unit in a structure has transacted. In specifications that include RC, RCI, Non-RC x RCI or RC x RCI interacted with Post, main effects of these variables are also included but not tabulated. Robust standard errors clustered by 1990 block group are in parentheses. See notes to Table 6 for additional details. *** p<0.01, ** p<0.05, * p<0.1

Table A6. Somerville Border Analysis
Dependent Variable: Log of Transacted Price

	All (1)	Houses (2)	Condominiums (3)
	I. Neighboring Census Tracts		
Cambridge Tract x Post	0.069***	0.066**	0.059
	(0.023)	(0.028)	(0.050)
Observations	5,700	4,398	1,302
R-squared	0.606	0.598	0.637
	II. Neighboring Census Block Groups		
Cambridge BG x Post	0.102**	0.101*	0.018
	(0.042)	(0.050)	(0.085)
Observations	2,775	1,991	784
R-squared	0.617	0.607	0.646

Sample includes transactions in Somerville that take place in a census tract or block group either abutting Cambridge (on the South) or Medford (on the North). Year fixed effects, property characteristics, and Census tract fixed effects included in all specifications. Robust standard errors clustered by 1990 block group are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Online Appendix

Data Appendix

Assessor's Data

The Cambridge Assessor's Database delineates the universe of residential housing located on each Cambridge "map-lot," which is Cambridge's internal land parcel numbering system. We assembled Assessor's data for 1995 and 2005, containing property valuations as of January 1 of the *prior* year (thus, we designate these files as "1994" and "2004" property assessments in the text). We obtained from the Cambridge Historical Commission and subsequently digitized bound copies of the 1995 Commitment Books, which contain the property type classification and assessed value of each Cambridge property, used for property tax purposes. We obtained a copy of the 2005 Assessor's Database in electronic form directly from the Cambridge Assessing Department. Unlike the 2005 data, the 1995 Commitment Books do not enumerate the number of units at each structure. In place of this enumeration, we use a file provided to us by Clifford Cook of the Cambridge Planning Department that contains the count of units in each structure at each map-lot in 2001. To calculate the latitude and longitude of each map-lot, we merged a geocoded version of the 2008 Assessor's Database provided by the MIT GIS Laboratory. We identified structure type conversions by comparing the structure types assessed at each map-lot in 1994 and 2004. The combined Assessor's files, augmented with structure counts and latitude and longitude data, comprise our residential structures file. For all assessment, transaction and investment data, we inflated nominal dollar values to 2008 dollars using the All Items Less Shelter CPI for All Urban Consumers, Series Id: CUUR0000SA0L2, Not Seasonally Adjusted, available from http://data.bls.gov/cgi-bin/srgate, last accessed May 2012.

Decennial Census Data

To determine the 1990 Census block, block group, and tract corresponding to each map-lot code, we used ESRI ArcMap and MassGIS ArcGIS shape-files containing Census geography boundaries, which allowed us to match map-lots to geographies by latitude and longitude. We obtained block group-level demographic data for the cities of Cambridge, Malden, Medford, and Somerville from the 1990 Census Summary Tape Files (STF-1 and STF-3), which enumerate detailed demographic and housing data by block group using either a 100 percent extract (STF-1) or 15 percent sample (STF-3) of the 1990 Census of Populations.

Rent Control Data

Our measure of rent control status and the geographic distribution of rent controlled properties is drawn from the (now defunct) Cambridge Rent Control Board's database of actively controlled properties as of 1994. This database was generously provided by David Sims of Brigham Young University, who obtained it from the City of Cambridge via an earlier Freedom of Information Act request. The Rent Control Board database lists the address and map-lot code of all structures that were actively rent controlled as of 1994. We designate a given map-lot as rent controlled for the purpose of our analysis if there are any actively rent controlled units on the map-lot as of 1994.

Sales Data

Data on transactions of houses and condominiums from 1988 and 2005 in Cambridge, Malden, Medford, and Somerville are from the residential real estate sales database, which was purchased from the Warren Group, a commercial vendor that assembles real estate data from town deeds offices. Since not all changes in ownership are conventional sales, we eliminated transactions that do not appear to be standard arms-length transactions, specifically: transactions where deeds are marked as coming from a foreclosure process or bearing a land court certification; transactions where the last name of the buyer or seller appears in the name of the party on the other side of the transaction; transactions involving the Cambridge Housing Authority or affordable housing entities such as Just-A-Start; properties where a seller and a buyer buy and sell the same property from each other on the same day; property share transactions, identified as an individual being on the same side of multiple transactions of the same property on the same day or the same individual being on both sides of any transaction; and transactions where the buyer resells the property later the same day. We also removed transactions with zero total rooms or zero interior square footage and retain only one copy of any duplicate transaction (those with the same street address, sale date, and price). In cases where the year built field for a given transacted property was missing, we first attempted to fill it in with the listed year built from other transactions for that property. For the two percent of transactions where year built could not be identified, we imputed its value as the mean year built for its structure type. In regressions that include property characteristics, we include a dummy variable equal to one if the year built value was imputed. We excluded 359 rent controlled condominium sales made between November 21, 1989 and December 31, 1989, during which time a portion of the rent control statute limiting condominium conversions and sales was temporarily overturned by the Massachusetts court.[49] To reduce the influence of outliers, we winsorized sales prices by structure type to the first percentile for the entire sample. Thus, transaction prices are defined as $\tilde{p}_i = \max\left\{p_i,\ p_{s(i),(.01)}\right\}$, where p_i is the (real) reported sales price of property i, and $p_{s(i),(.01)}$ is the first percentile of housing sale prices for structure type $s(i)$.

Investment Data

To measure residential investments, we obtained from the Cambridge Inspectional Services Department a database of all residential building permits issued by the City of Cambridge between 1991 and 2005. For 389 permits missing expenditure amounts but containing information on the permit fee, we replaced the missing expenditure value with 100 times the 1% permit fee. We removed duplicate permits and permits that were not designated for residential properties, specifically those that mention a non-residential usage in the description field (e.g., business, office, tent, educational, store, mixed, commercial, research, etc.). Since the investment permit data do not contain map-lot codes, we pooled the Assessor's Database with the Cambridge Rent Control Board database and the 2001 structures file to form a crosswalk between address strings and map-lot codes. For permits lacking an address string that matched to a map-lot code, we matched the permit to the nearest map-lot code based on straight-line distance, calculated using the StreetMaps USA address locator in ArcGIS. To construct our investment analysis sample, we merged the investment data to the housing structures file using map-lot codes to determine rent control status, proximity to rent control, property type, and geographic location. We summed all permitted expenditures at a map-lot in a year to form an annual panel of residential map-lot codes containing total expenditure for each map-lot and winsorized real investment expenditures to the 99.5$^{\text{th}}$ percentile for each structure type and year.

[49]See Massachusetts Supreme Judicial Court *406 Mass. 147*, detailed at http://masscases.com/cases/sjc/406/406mass147.html, last accessed May 2012. The City of Cambridge was able to quickly revise the rent control statute to comply with the court ruling while again limiting the conversion and sale of rent controlled apartments.

Cambridge City Census Data

To build a longitudinal panel database of all adult Cambridge residents for the years 1991 through 2000, we digitized the Cambridge City Census files from 1991 through 2000, obtained from the Cambridge Election Office, to form a comprehensive list of adult Cambridge residents containing for each resident full name, address, birth year, occupation, country of birth, and optionally, gender and political party. Cambridge collects and makes publicly available these data in accordance with Massachusetts law requiring each municipality to conduct an annual census for purposes of voter registration and state reimbursements. We contracted with Equifax, Inc., a major U.S. credit bureau, to match the names and addresses of the approximately 436,000 adult Cambridge residents (39,000 to 48,000 unique individuals per year) identified in the city census. Equifax provided a unique identification number for each queried Cambridge resident, which allowed us to link residents across years and addresses to identify individuals who remained at the same address in consecutive years. Because the City Census files do not contain map-lot codes, we matched the address of each resident to its map-lot using the crosswalk between map-lot code and address strings constructed for the investment analysis sample.

Abt Associates Survey Data

We analyzed rent differentials at controlled relative to non-controlled units using data from an Abt Associates study (Finkel and Wallace, 1987) commissioned by the City of Cambridge to gather data on the characteristics of households living in rent controlled housing. These data, provided to us by Clifford Cook of the Cambridge Planning Department, enumerate contract rent, rent control status, tenant awareness of rent control status, unit characteristics (bedrooms, bathrooms, total rooms, an indicator for elevator in building, and indicators for whether furnishings, heat, electricity, or water were included in the rent), zip code, and variables indicating unit condition for a sample of 906 units.

Merging Assessor, Transaction, Investment and Geographic Data

To form the analytic sample for assessments and transactions, we mapped the housing structures file by latitude and longitude to 1990 U.S. Census geographies (tracts, block groups, and blocks) using a spatial join in ArcGIS. We merged the combined database to the Rent Control Board file according to map-lot code or, where necessary, street address. We calculated Rent Control Intensity (RCI) at each map-lot as the fraction of housing units within a given radius, according to longitude and latitude, that were rent controlled (excluding the map-lot's own rent control status from the calculation).

To pair RCI information with transactions data, we merged the Warren Group data with the housing structures file by map-lot code. For transactions with a missing map-lot code or with a map-lot code that did not merge with the structures file, we queried the street address in Cambridge's online property database (http://www2.cambridgema.gov/fiscalaffairs/propertysearch.cfm) and recorded the online entry's map-lot code.

Additional Empirical Results

Price appreciation in adjoining cities

The decade following the elimination of rent control in Cambridge saw substantial housing price appreciation throughout Massachusetts. For example, the Federal Housing Finance Agency's OFHEO house price index (HPI) of single-family houses for the Boston Metropolitan Statistical Area (MSA), an area corresponding to 97 towns including Cambridge (which accounts for 3% of the total MSA population), shows a 270 percent increase from the first quarter of 1995 to the first quarter of 2005.[50] This backdrop of rising real estate prices raises a potential concern that the price appreciation in Cambridge that we attribute to rent decontrol might instead reflect aggregate house price trends. Since our empirical strategy compares price appreciation across local areas *within* Cambridge, this aggregate phenomenon is only a threat to our identification strategy if it leads to *differential* appreciation at formerly rent control-intensive locations for reasons that are unrelated to rent decontrol.

One way to explore this concern is to compare price appreciation in rent control-intensive locations in Cambridge to comparable locations in surrounding Massachusetts towns that did not have rent control in this time period. We implement this comparison by analyzing housing transaction data for the three nearby cities: the adjoining city of Somerville, which abuts Cambridge; the city of Medford, which abuts Somerville; and the city of Malden, which abuts Medford. These transactions

[50] Genesove and Mayer (2001) also document price appreciation in downtown Boston's condominium market during the 1990s.

data are also sourced from the Warren Group files, used for the price analysis immediately above, and contain the identical data elements and years of coverage.[51]

To perform the comparison, we create a Predicted RCI ('P-RCI') measure for Cambridge and surrounding towns, first by regressing the Cambridge block group level RCI measure on 18 distinct block group attributes available from the 1990 Census Summary Tape Files (STF) to obtain a forecasting relationship between census block group attributes and RCI in Cambridge.[52] We next use the model to predict the P-RCI value for each block group in Cambridge, Somerville, Malden, and Medford. Finally, we explore the relationship between P-RCI and residential real estate price appreciation within all four cities. Table B3 presents estimates.

As a benchmark, the first pair of models in Panel I presents the relationship between *actual* RCI and pre-post decontrol price appreciation at never-controlled house and condominium properties in Cambridge between 1988 and 2005.[53] The RCI measure in this specification is computed at the 1990 Census block group level. Panel II presents an identical set of estimates where the observed RCI measure is replaced by P-RCI. For Cambridge houses, the point estimate for the the Post $\times$ P-RCI variable is reassuringly similar to the estimate using the actual RCI measure. This suggests that we may be able to use the statistical relationship between RCI and block group census attributes to construct a proxy for RCI in non-Cambridge towns. The correspondence is not as close for Cambridge condominiums, however, which limits the informativeness of this exercise for this property type.

The next four panels of Table B3 perform the comparison exercise using transaction data from Somerville, Malden, and Medford. Panel III pools these three cities. Distinct from the pattern for Cambridge houses, we detect neither a significant negative relationship between P-RCI and house

[51]A caveat to this approach is that it is widely perceived in the Cambridge area (though nowhere documented to our knowledge) that the displacement of Cambridge residents following decontrol—both those leaving decontrolled units and those fleeing rising rents—spurred gentrification of parts of Medford and Somerville. Lending some credence to this hypothesis, Atlantic Marketing Research (1998) reports that 58 percent of Cambridge renters who moved out of their decontrolled units between 1994 and 1997 left Cambridge. In general, we would expect this potential spillover to surrounding cities to bias us towards finding similar differential rises in property prices in non-Cambridge comparison neighborhoods.

[52]The 18 block group attributes are population density, median family income, the fraction of commuters using public transportation, average owner tenure (the average tenure in years of owner-occupants at their current residence), average renter tenure, the fraction of owner-occupied housing units that were built before 1970, the fraction of renter-occupied units built before 1970, the fraction of units that are condominiums, the fraction of residents that are renters, the number of residents within non-family households (e.g. roommates), average age, median contract rent, the average residential property value, the fraction of residents self-identifying as white, the fraction of residents self-identifying as Asian, the fraction of housing units that are vacant, the fraction of housing units in structures with at least 20 units, and the fraction of housing units in structures with 5 to 19 units

[53]We exclude decontrolled properties from this exercise since no such properties exist in the comparison cities.

transaction prices in the pre-decontrol period nor a significant positive relationship between P-RCI and house transaction prices in the post-decontrol period. Panels IV, V, and VI report estimates separately for the three non-Cambridge cities. These models find inconsistently signed relationships between P-RCI and house transaction prices in Somerville, Malden, and Medford. One surprising result, however, is that the point estimate for P-RCI $\times$ Post for Medford houses is similar to the analogous estimate for Cambridge houses and is significant at the 5 percent level—though unlike in Cambridge, the relationship between P-RCI and house transaction prices in Medford is small and statistically insignificant in the pre-decontrol period. While this result is disconcerting inasmuch as Medford did not have rent control, we are inclined to view this as a chance finding given the evidence in the prior three panels.

A second pattern in Table B3 is that for all towns *except* Cambridge, we find evidence of a substantial decline in the transaction prices of condominiums in block groups with high P-RCI in the post-decontrol period. Placing this result in context, it bears note that Massachusetts experienced a substantial increase in condominium construction and conversions in urban neighborhoods in this period, and this supply shift may have lowered prices. Comparing the non-Cambridge condominium results to those for Cambridge, one potential inference is that condominium prices in rent control-intensive neighborhoods in Cambridge would have fallen substantially after 1994 had it *not* been for the end of rent control. Given the many complexities surrounding condominium supplies and prices in this time period, however, we remain agnostic on this point.

Changes in the Composition of Properties Transacted

Because not all properties transact before or after the end of rent control, a concern in interpreting models involving the prices of transacted properties is that non-random selection of properties into transaction could lead to biased estimates of the causal effects of interest. To examine selection into transaction, we estimated Seemingly Unrelated Regression (SUR) models that explore how the characteristics of transacted properties vary with rent control status and rent control intensity in Tables B4 and B5.[54] For houses, shown in Table B4, we detect no individually or jointly significant pre-post decontrol changes in the relationship between rent control status, rent control intensity, and selection into transaction. There are some compositional differences for condominiums, however,

[54]The regression specification for these models is comparable to those above, except in place of house values, we use as dependent variables a vector of property characteristics and the full set of equations is fit simultaneously to allow for hypothesis testing across equations.

shown in Table B5. Condominiums that transact after 1994 at decontrolled locations have 0.15 fewer total rooms than those that transact prior to 1995 and they are more likely to be recently built. Moreover, the number of total rooms and bedrooms in transacted decontrolled condominiums increases after 1994 in rent control-intensive neighborhoods, a pattern consistent with upgrading of these locations. The chi-squared test for the joint significance of these relationships confirms that the end of rent control saw a shift in the composition of transacted condominiums, which suggests some caution in drawing conclusions using the transaction price data for condominiums.

As a second method of addressing potential composition biases in the transaction sample, in Table B6, we re-estimated the Table 7 models for price appreciation using only the subset of units that did not change structure type between 1994 and 2004 (typically reflecting condominium conversions), thus reducing the samples of house and condominium transactions by 6 and 21 percent respectively. This sample restriction reduces the estimated indirect effect of decontrol on never-controlled units by approximately 25 percent and on decontrolled units by approximately one-half. The greater effect on decontrolled units in turn reflects the fact that a substantial number were converted to condominiums and renovated, which likely contributed to their price appreciation.

Availability of Subprime Credit

To examine the possibility that the appreciation of Cambridge properties in formerly rent control intensive neighborhoods was driven in part by the availability of subprime lending into lower income neighborhoods (Mian and Sufi, 2009), in the bottom panel of Table B6, we again re-estimated the Table 7 specifications while excluding transactions in which the mortgage lender is listed as a subprime lender by the U.S. Department of Housing and Urban Development.[55] As it turns out, only two percent (311) of our sample of 14,789 transacted properties in Cambridge were financed by subprime lenders, and hence excluding these observations has no substantive effect on the findings.

[55] This follows the approach of Gerardi, Shapiro and Willen (2007a), who use U.S. Department of Housing and Urban Development data to construct a list of subprime lenders.

Table B1. Effects of Rent Decontrol on Assessed Values with Alternative Block Group Control Strategies

Dependent Variable: Log of Assessed Property Value (1994, 2004)

	(1)	(2)	(3)
		I. All	
RC x Post	0.220***	0.196***	0.200***
	(0.036)	(0.035)	(0.037)
Non-RC x RCI x Post	0.514**	0.418**	0.465**
	(0.227)	(0.209)	(0.212)
RC x RCI x Post	0.651***	0.674***	0.734***
	(0.231)	(0.226)	(0.239)
N	32,980	32,779	32,779
R-squared	0.767	0.846	0.849
		II. Houses	
RC x Post	0.060***	0.061***	0.057***
	(0.011)	(0.012)	(0.011)
Non-RC x RCI x Post	0.194*	0.208*	0.193*
	(0.103)	(0.109)	(0.106)
RC x RCI x Post	0.315**	0.247*	0.265*
	(0.130)	(0.133)	(0.134)
N	16,239	16,155	16,155
R-squared	0.855	0.902	0.917
		III. Condominiums	
RC x Post	0.350***	0.329***	0.332***
	(0.040)	(0.039)	(0.040)
Non-RC x RCI x Post	0.678**	0.369	0.395
	(0.308)	(0.260)	(0.271)
RC x RCI x Post	0.648**	0.603**	0.687**
	(0.291)	(0.268)	(0.261)
N	16,741	16,624	16,624
R-squared	0.714	0.836	0.828
Fixed Effects	Block Group	Rolling	Exponential Decay

RCI is calculated over a 0.20 mile radius and demeaned. RC is an indicator for a location that was rent controlled in 1994 and Post is an indicator for year equal to 2004. Year fixed effects and structure type dummies are included in all regressions. All specifications include tract trends, which are tract x Post dummies for each of 30 tracts from the 1990 Census. Block group fixed effects correspond to the 89 Cambridge block groups in the 1990 Census containing assessed properties. "Rolling" block fixed effects assign units to non-mutually exclusive Census block dummies by identifying each Census block whose centroid lies within 0.2 miles of a residential map-lot and assigning these block dummies to the map-lot. "Exponential Decay" specifications are block fixed effects based on an exponential decay (with lambda=1), where all block dummies are fractionally assigned (summing to 1) to each map-lot as a decaying function of the distance between the property's map-lot and the block centroid. Robust standard errors clustered by 1990 block group are in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table B2. Effects of Rent Decontrol and Rent Control Intensity on Assessed Values by Structure Type Dropping Block Groups Bordering non-Cambridge Properties

	(1)	(2)	(3)	(4)	(5)	(6)
			I. Houses			
RC x Post	0.091***	0.073***	0.083***	0.064***	0.080***	0.076***
	(0.016)	(0.013)	(0.015)	(0.012)	(0.016)	(0.015)
RCI x Post	0.420***	0.317***			0.434***	
	(0.088)	(0.094)			(0.085)	
Non-RC x RCI x Post			0.403***	0.297***		0.423***
			(0.088)	(0.094)		(0.084)
RC x RCI x Post			0.597***	0.525***		0.556***
			(0.153)	(0.120)		(0.159)
H_0: No Spillovers			0.00	0.00		0.00
H_0: Spillovers Equal	-	-	0.122	0.00305	-	0.309
R-squared	0.847	0.850	0.847	0.850	0.850	0.850
N	12,407	12,407	12,407	12,407	11,415	11,415
			II. Condominiums			
RC x Post	0.335***	0.343***	0.336***	0.335***	0.322***	0.324***
	(0.050)	(0.041)	(0.053)	(0.041)	(0.046)	(0.045)
RCI x Post	0.230	0.691**			0.101	
	(0.170)	(0.282)			(0.196)	
Non-RC x RCI x Post			0.283	0.571**		-0.040
			(0.185)	(0.275)		(0.126)
RC x RCI x Post			0.123	0.744**		0.213
			(0.321)	(0.325)		(0.357)
H_0: No Spillovers			0.289	0.0415		0.818
H_0: Spillovers Equal			0.669	0.593		0.528
R-squared	0.701	0.710	0.701	0.710	0.712	0.712
N	14,140	14,140	14,140	14,140	10,131	10,131
Block group FEs	y	y	y	y	y	y
Tract Trends	-	y	-	y	-	-
Excluding Converted	-	-	-	-	y	y

Notes. Dependent variable is log assessed value. Assessed values are from 1994 and 2004. RCI is calculated based on a radius of 0.20 miles. RCI is demeaned so that the main rent control effect is the price differential for the average rent controlled property. Post is an indicator for year equal to 2004. Year fixed effects and block group fixed effects are included in all regressions. Block group fixed effects correspond to each of the 88 Cambridge block groups containing assessed properties using 1990 Census boundaries. Tract trends are tract x post dummies for each of 30 tracts from the 1990 Census. Two-family and three-family house dummies are included in the top panel. Test for No Spillovers report p-vaues from tests that RCI x Post or Non-RC x RCI x Post and RC x RCI x Post coefficients are zero. Test for Spillovers Equal reports p-value for hypothesis that these latter two coefficients are equal. Robust standard errors clustered by 1990 block group are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B3. Placebo Estimates of the Relationship between Imputed Rent Control Intensity and Property Price Appreciation in Cambridge and Adjoining Cities, 1988 - 2005

Dependent Variable: Log Sale Price

	Houses	Condo-miniums	Houses	Condo-miniums	House	Condo-miniums
	(1)	(2)	(3)	(4)	(5)	(6)
	I. Cambridge: Actual RCI		II. Cambridge: Predicted RCI		III. Somerville, Medford and Malden	
RCI	-0.183	-0.257	-0.203**	-0.504*	-0.034	0.101
	(0.112)	(0.226)	(0.096)	(0.256)	(0.057)	(0.205)
RCI x Post	0.261***	0.063	0.278***	-0.055	0.088	-0.574***
	(0.088)	(0.093)	(0.092)	(0.102)	(0.055)	(0.206)
N	4,223	5,764	4,223	5,764	17,270	3,346
	IV. Somerville		V. Malden		VI. Medford	
RCI	-0.162	0.238	0.023	-0.176	-0.056	0.832***
	(0.133)	(0.555)	(0.077)	(0.172)	(0.079)	(0.268)
RCI x Post	-0.090	-0.406	0.052	-0.562***	0.174**	-1.201***
	(0.151)	(0.507)	(0.066)	(0.171)	(0.086)	(0.278)
N	6,605	1,868	6,506	1,197	4,159	281

Sample includes never-controlled houses and condominiums in Cambridge and surrounding cities transacted during 1988 through 2005. Prices are winsorized by structure type and city at the first percentile. Actual RCI is Rent Control Intensity calculated at the 1990 Census block group level. Predicted RCI is imputed for Cambridge and non-Cambridge block groups from an OLS regression of Cambridge block group RCI on 1990 Cambridge Census block group characteristics. All specifications include the controls used in column 1 of Table 7. Panel III additionally includes a full set of city by year effects. Robust standard errors clustered by 1990 block group are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B4. Seemingly Unrelated Regression Estimates for Changes in Attributes of Transacted Houses Following Rent Control Removal

	Total Rooms (1)	Bathrooms (2)	Bedrooms (3)	Interior Sqft (10s) (4)	Lot Size Sqft (100s) (5)	ln(Age) (6)	χ^2 Test (row) (7)
			I. Models with Common RCI Effect				
Constant	7.26***	2.46***	3.16***	204.17***	23.40***	4.83***	
	(0.36)	(0.12)	(0.22)	(13.20)	(4.19)	(0.11)	
RC x Post	-0.16	-0.05	-0.00	1.53	1.62	-0.09	6.44
	(0.20)	(0.06)	(0.12)	(7.04)	(2.38)	(0.06)	(0.38)
RCI x Post	0.20	0.03	0.02	18.09	-0.44	0.04	3.13
	(0.46)	(0.14)	(0.28)	(15.83)	(5.35)	(0.13)	(0.79)
			II. Models with RC x RCI Interactions				
Constant	8.10***	2.46***	3.17***	204.17***	25.91***	4.76***	
	(0.38)	(0.12)	(0.22)	(13.20)	(4.46)	(0.10)	
RC x Post	-0.09	-0.04	0.03	2.87	2.09	-0.09	6.04
	(0.21)	(0.07)	(0.13)	(7.27)	(2.46)	(0.06)	(0.42)
Non-RC x RCI x Post	0.46	0.06	0.19	22.36	1.54	0.03	4.22
	(0.48)	(0.15)	(0.30)	(16.70)	(5.64)	(0.14)	(0.65)
RC x RCI x Post	-1.92	-0.24	-1.14	-18.13	-15.26	0.11	2.43
	(1.43)	(0.45)	(0.88)	(49.45)	(16.70)	(0.41)	(0.88)
H_0: No Spillovers	0.257	0.792	0.350	0.380	0.634	0.939	
H_0: Spillovers Equal	0.114	0.518	0.151	0.437	0.340	0.856	

N = 4,814. Table reports estimates from Seemingly Unrelated Regressions of characteristics of transacted houses on 1994 rent control status and Rent Control Intensity (RCI) calculated at the 0.20 mile radius. All specifications include main effects for RC, RCI or Non-RC x RCI and RC x RCI, year of sale dummies, structure type dummies, block-group fixed effects, and an indicator for whether year built was imputed. Column 7 reports Chi2(6) tests for the null hypothesis that the given row's coefficients are jointly equal to zero (with p-values in parentheses). Test for No Spillovers reports p-values from tests that RCI x Post or Non-RC x RCI x Post and RC x RCI x Post coefficients are zero. Test for Spillovers Equal reports p-values from tests that these latter two coefficients are equal. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B5. Seemingly Unrelated Regression Estimates for Changes in Attributes of Transacted Condominiums Following Rent Control Removal

	Total Rooms (1)	Bathrooms (2)	Bedrooms (3)	Interior Sqft (10s) (4)	Has Lot (5)	ln(Age) (6)	χ^2 Test (row) (7)
				I. Models with Common RCI Effect			
Constant	3.41***	1.50***	1.43***	91.64***	1.05***	2.01***	
	(0.17)	(0.07)	(0.08)	(7.34)	(0.02)	(0.09)	
RC x Post	-0.15**	0.03	-0.03	-2.67	0.02***	-0.55***	186.46
	(0.07)	(0.03)	(0.04)	(2.95)	(0.01)	(0.04)	(0.00)
RCI x Post	0.04	-0.19**	0.04	-4.50	-0.00	0.09	9.77
	(0.22)	(0.08)	(0.11)	(9.16)	(0.02)	(0.13)	(0.13)
				II. Models with RC x RCI Interactions			
Constant	3.40***	1.59***	1.49***	97.19***	1.04***	2.51***	
	(0.16)	(0.06)	(0.09)	(6.66)	(0.02)	(0.11)	
RC x Post	-0.19***	0.03	-0.05	-3.46	0.02**	-0.55***	180.96
	(0.07)	(0.03)	(0.04)	(3.02)	(0.01)	(0.04)	(0.00)
Non-RC x RCI x Post	-0.28	-0.17	-0.12	-9.62	0.01	0.02	17.86
	(0.26)	(0.10)	(0.13)	(11.09)	(0.03)	(0.15)	(0.01)
RC x RCI x Post	0.71*	-0.24	0.40**	6.54	-0.02	0.25	2.70
	(0.38)	(0.15)	(0.20)	(16.19)	(0.04)	(0.22)	(0.85)
H_0: No Spillovers	0.101	0.072	0.082	0.633	0.873	0.531	
H_0: Spillovers Equal	0.033	0.668	0.028	0.410	0.605	0.399	

N=9,975. Table reports estimates from Seemingly Unrelated Regressions of characteristics of transacted condominiums on 1994 rent control status and Rent Control Intensity (RCI) calculated at the 0.20 mile radius. All specifications include main effects for RC, RCI or Non-RC x RCI and RC x RCI, year of sale dummies, structure type dummies, block-group fixed effects, and an indicator for whether year built was imputed. Has Lot is a dummy variable for whether the transacted condominium had an accompanying lot. Column 7 reports Chi2(6) tests for the null hypothesis that the given row's coefficients are jointly equal to zero (with p-values in parentheses). Test for No Spillovers reports p-values from tests that RCI x Post or Non-RC x RCI x Post and RC x RCI x Post coefficients are zero. Test for Spillovers Equal reports p-values from tests that these latter two coefficients are equal. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B6. Relationship between Rent Control, Rent Control Intensity and Transaction Price, 1988 - 2005: Eliminating Units that Were Converted from their 1994 Structure Type and Transactions Financed by Subprime Lenders

Dependent Varaiable: Log Sale Price

	All transactions			Houses			Condominiums		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				I. Eliminating Converted Structures					
RC x Post	0.072***	0.099	0.061	0.093**	0.104**	0.106**	0.076**	0.072**	0.065**
	(0.027)	(0.069)	(0.072)	(0.045)	(0.046)	(0.047)	(0.031)	(0.028)	(0.027)
RCI x Post	0.145***			0.361***			0.029		
	(0.055)			(0.082)			(0.072)		
Non-RC x RCI x Post		0.155**	0.069		0.389***	0.290**		-0.025	-0.153
		(0.070)	(0.102)		(0.091)	(0.136)		(0.107)	(0.146)
RC x RCI x Post		0.093	0.085		0.054	-0.034		0.149	0.195
		(0.150)	(0.174)		(0.282)	(0.297)		(0.154)	(0.180)
N	12,402	12,402	12,402	4,527	4,527	4,527	7,875	7,875	7,875
				II. Eliminating Subprime Lenders					
RC x Post	0.088***	0.057	0.023	0.085**	0.096**	0.097**	0.095***	0.081***	0.071**
	(0.026)	(0.068)	(0.067)	(0.041)	(0.042)	(0.043)	(0.030)	(0.029)	(0.027)
RCI x Post	0.204***			0.339***			0.152**		
	(0.058)			(0.079)			(0.073)		
Non-RC x RCI x Post		0.191***	0.121		0.361***	0.268**		0.074	-0.014
		(0.068)	(0.091)		(0.086)	(0.126)		(0.086)	(0.130)
RC x RCI x Post		0.256*	0.249		0.092	-0.013		0.317**	0.307
		(0.146)	(0.177)		(0.253)	(0.282)		(0.154)	(0.197)
N	14,478	14,478	14,478	4,706	4,706	4,706	9,772	9,772	9,772
Quadratic Tract Trends	-	-	y	-	-	y	-	-	y

Table reports estimates for all residential transactions, and houses and condominums separately following Table 7. Panel I drops transactions of units in structures converted from their 1994 structure type. Panel II drops transactions with mortgages from lenders identified by HUD as issuing subprime mortgages. Robust standard errors clustered by 1990 block group are in parentheses. See Table 7 notes for additional details. *** p<0.01, ** p<0.05, * p<0.1

Chapter 3

IV Quantile Regression for Group-level Treatments, with an Application to the Effects of Trade on the Distribution of Wages

with Denis Chetverikov and Bradley Larsen

3.1 Introduction

In classical linear panel models, when time-invariant unobservables are correlated with included variables, unit fixed effects are commonly used to obtain identification. While this approach yields consistent estimates of coefficients on time-varying characteristics, it rules out the identification of effects of any time-invariant variables, as these variables are eliminated by the within-group transformation. In an influential paper, Hausman and Taylor (1981) demonstrated that within variation can be used to identify the coefficients on time-dependent (individual-level) covariates, and then exogenous between variation can be used to identify coefficients of any time-invariant (group-level) covariates.[1] Our paper provides a robust method of estimating the effect of group-level endogenous

[1] To clarify the broad applicability of the estimator to a variety of settings, we depart from the usual panel-data terminology and refer to panel units as groups (instead of as individuals) and to within-group observations as individuals or micro-level observations (instead of as time observations).

covariates on the distribution of outcomes, a quantile extension of the Hausman and Taylor (1981) classical linear panel estimator. We demonstrate that standard quantile regression techniques, such as Koenker and Bassett (1978), are inconsistent in this setting, and that existing instrumental variables methods for quantile regression do not apply. We find that our estimator is computationally simple to implement and, when group-level unobservables are present, substantially outperforms traditional methods for obtaining distributional effects of a group-level treatment.

Our paper makes the following theoretical and practical contributions. First, we show how to use the Bahadur representation of quantile estimators to derive weak conditions on the growth of the number of observations per group that are sufficient for the consistency and (mean-zero) asymptotic normality of our estimator. Note that theoretical results for the classical linear panel model do not require these tools, as the linear panel model admits the linear fixed effect estimator, which is unbiased. This feature of linearity allowed Hausman and Taylor (1981) to derive a consistent estimator of coefficients corresponding to the group-level endogenous covariates using asymptotics where the number of observations per group is fixed. In contrast, quantile estimators are generally biased in finite samples, raising challenges for obtaining a consistent estimator of these coefficients in a quantile panel data model. However, quantile estimators are asymptotically unbiased, motivating us to adopt the Bahadur representation to obtain our theoretical results.

The estimator we propose is of practical significance in situations in which the researcher has data on a group-level endogenous treatment and has micro-level data on the outcome of interest within each group. For example, a researcher may be interested in a policy which varies across states and wish to examine the effects of the policy on the within-state distribution of individual outcomes. The estimator consists of 1) performing quantile regression within each group to estimate effects of micro-level covariates, or, if no micro-level covariates are included, calculating the desired quantile for the outcome within each group; and 2) applying two-stage least squares (2SLS) to the group-level quantiles. Importantly, as in Hausman and Taylor (1981), micro-level covariates can be used as internal instruments for the endogenous group-level treatment if they satisfy the exogeneity condition. Alternative approaches for IV quantile regression, in which the researcher is concerned with an individual-level treatment being correlated with unobservables, such as the situations described in Abadie, Angrist and Imbens (2002) or Chernozhukov and Hansen (2005), do not apply in this setting. Angrist and Pischke (2009) described IV quantile methods as "a relatively new development and not yet as flexible as conventional 2SLS." An advantage of the estimator proposed here is that it is as flexible as 2SLS.

Even in the absence of endogeneity, standard quantile regression will be inconsistent in this setting, as demonstrated below, given that an additive unobservable component breaks the quantile structure. Our approach, on the other hand, consists of a linear second stage, in which the additive unobservable component is not problematic. Intuitively, when group-level variables are exogenous and there are no individual-level covariates—a special case of our general model—the second-stage estimator reduces to a between estimator with group quantiles replacing group means, equivalent to the minimum distance estimator proposed in Chamberlain (1994) and discussed further in Angrist, Chernozhukov and Fernández-Val (2006). It is also useful to note that the additive unobservables in our model, when uncorrelated with group-level included variables, are akin to left-hand side measurement error. While posing no problems for linear models, left-hand side errors-in-variables can severely bias quantile estimation, as we demonstrate below.[2] To our knowledge, this use of the Chamberlain (1994) estimator to solve the bias introduced by group-level unobservables in quantile models (or, equivalently, left-hand side measurement error), while still allowing for the estimation of group-level treatments, has not before been discussed.

As in Hausman and Taylor (1981), estimating coefficients corresponding to group-level covariates requires estimating coefficients corresponding to micro-level covariates in a first stage. Estimating the effects of micro-level covariates in classical panel data models can be easily accomplished by taking differences or by demeaning observations within groups. This trick, however, does not work in the quantile panel data model because it breaks the quantile structure. A possible alternative in the panel data model is to jointly estimate group-level effects together with the effects of micro-level covariates. This procedure is computationally burdensome, however, because the number of parameters in the quantile optimization problem in this case is at least as big as the number of groups in the data set, which is typically large.[3] We find, in Monte Carlo experiments, that jointly estimating group-level effects takes over 130 times as long as our approach.[4]

[2]Hausman (2001), footnote 11, also pointed out that left-hand side measurement error will bias quantile regression estimates. See Hausman, Luo and Palmer (2014) for an EM algorithm-based solution to left-hand side measurement error in quantile models.

[3]Koenker (2004), Lamarche (2010), Galvao and Wang (2013), and Galvao (2011) all highlighted the computational challenge present when estimating large numbers of fixed effects in quantile models. Koenker (2004) explained how sparse matrix approaches can be used to aid computation in some cases.

[4]Note that, even when no group-level unobservables are present—meaning standard quantile regression is consistent in addition to grouped quantile regression being consistent—the grouped quantile regression approach still provides computational advantages over standard quantile regression. A prime example is a model with state fixed effects where the group is a state-by-year combination, as in many policy analyses; grouped quantile regression ignores state fixed effects in the first stage and then accounts for them linearly in the second stage, greatly reducing the num-

A final contribution of practical significance is that standard errors for the proposed estimator can be obtained using traditional robust variance estimators for 2SLS, making inference particularly simple. We demonstrate that the estimator is consistent and asymptotically normal. We also show the existence of confidence bands for the coefficient of interest which hold uniformly over a set of quantiles of interest, and present subsampling and bootstrap approaches for estimating these confidence bands.

Section 3.2 presents the model and discusses its relationship to existing models. Section 3.3 provides several motivating examples of settings in which this estimation approach can be useful, drawing in part on papers which applied special cases of our approach but which did not discuss the consistency of the estimator. The first example comes from Angrist and Lang (2004), who estimated the effect on the lower tail of student test scores when urban students were reassigned to suburban schools through Boston's Metco program. The second example is Larsen (2012b), who estimated the effect of teacher certification laws on the distribution of teacher quality. The third example comes from Palmer (2011), and illustrates the use of our estimator in an IV setting, measuring the effects of suburbanization on urban decline. The final example suggests a new test of the symmetric, independent private values assumption in English auctions which could be used in a non-IV setting or in an IV setting given appropriate instruments for the number of bidders.

In Section 3.4 we discuss the estimator in detail. We examine the general case in which the researcher has individual-level outcome data and a treatment that varies by group, where the treatment is correlated with an additive, unobservable, group-level element. We describe the steps to the general grouped IV quantile estimator and its special cases. We then provide an example demonstrating that standard quantile regression is inconsistent in this setting.

Section 3.5 derives the asymptotic properties of the estimator. Our theoretical results are based on the asymptotics where both the number of groups and the number of observations per group grow to infinity. In contrast with other papers studying quantile panel data models, our results on asymptotic (mean-zero) normality hold under a weak assumption that the number of observations per group grows at least as quickly as the square-root of the number of groups (up to logarithmic terms). Other papers require at least linear growth; see, for example, Kato and Galvao (2011) and Kato, Galvao and Montes-Rojas (2011). This gives an advantage when the number of observations per group is relatively small in comparison with the number of groups.

ber of parameters to be estimated in the quantile model and hence reducing the computational burden significantly. See Section 3.3 for further discussion.

We present Monte Carlo simulations in Section 3.6 demonstrating the performance of our estimator. Findings indicate that, when the treatment variable is endogenous, group-level unobservables lead to bias in standard quantile regression. The grouped IV quantile estimator, on the other hand, has very low bias. The simulations also demonstrate that standard quantile regression is biased even if the treatment variable is exogenous, as group-level unobservables bias quantile regression just as left-hand side measurement error would. The grouped approach derived in this paper again has very low bias in this setting. We show that these results are true even with few groups and few individual observations per group.

Section 3.7 presents our main empirical application, which studies the effect of trade, in the form of increased import competition from China, on the distribution of wages within US local labor markets. We build on the work of Autor, Dorn and Hanson (2013), who studied the effects of Chinese import competition on average wages in local labor markets. In this setting, if there exists an unobservable component to wages in each local labor market, standard quantile regression would be inconsistent, and existing instrumental variables approaches to quantile regression are not applicable. Using the grouped IV quantile regression approach, we find that Chinese import competition harmed low-wage earners more than high-wage earners, particularly for females, an effect which is overlooked by traditional 2SLS.

To the best of our knowledge, our paper is the first to present a framework for estimating distributional effects as a function of group-level covariates. There is, however, a large literature studying quantile models for panel data when the researcher wishes to estimate distributional effects of individual-level covariates. See, for example, Koenker (2004), Abrevaya and Dahl (2008), Lamarche (2010), Canay (2011), Galvao (2011), Kato et al. (2011), Kato and Galvao (2011), Galvao and Wang (2013), and Arellano and Bonhomme (2013). Our paper also contributes to the growing literature on IV treatment effects in quantile models, such as Abadie et al. (2002), Chernozhukov and Hansen (2005), Lee (2007), Chesher (2003), and Imbens and Newey (2009). Our paper differs, however, in that this literature focuses on estimating distributional effects when the treatment is correlated with an individual-level quantile rather than focusing on group-level, additive unobservables.

We will use the following notation. Let c and C denote generic strictly positive constants that are independent of n but may vary at each appearance. Let $\|\cdot\|$ denote the Euclidean norm. Let $B(x, r)$ denote a ball with the center at x and radius r in a metric space that should be clear from the context. The symbol $\Rightarrow$ signifies weak convergence, and $l^\infty(\mathcal{U})$ for a set $\mathcal{U}$ denotes a set of bounded functions on $\mathcal{U}$. All proofs are contained in the Appendix.

3.2 Model

We study the following model for the response variable y_{ig} of individual i in group g,

$$y_{ig} = z_{ig}'\gamma(u_{ig}) + x_g'\beta(u_{ig}) + \varepsilon(u_{ig}, \eta_g), \tag{3.1}$$

where z_{ig} and x_g are d_z- and d_x-vectors of individual and group-level observable covariates (x_g contains the constant), η_g is a vector of group-level unobservable covariates, and u_{ig} is a scalar random variable representing individual heterogeneity. To turn this model into a quantile regression, we assume that for any given value of (z_{ig}, x_g, η_g) on its domain, y_{ig} is increasing in u_{ig}, and u_{ig} is distributed uniformly on $[0,1]$. Given these assumptions, the conditional uth quantile of y_{ig} given (z_{ig}, x_g, η_g) is

$$Q_{y_{ig}|z_{ig},x_g,\eta_g}(u) = z_{ig}'\gamma(u) + x_g'\beta(u) + \varepsilon(u, \eta_g).$$

We assume that a researcher has data on G groups with N_g individuals within group g, $g = 1, ..., G$. In this paper, we are interested in estimating and performing inference on the parameter $\beta(u)$ representing the effect of observable group-level covariates either for some particular value of index u or for a set of indices $\mathcal{U}$.

In this model, $z_{ig}'\gamma(u_{ig})$ is the individual effect, and $x_g'\beta(u_{ig}) + \varepsilon(u_{ig}, \eta_g)$ is the group effect. We allow for arbitrary correlation between these effects. The unobservable component $\varepsilon(u_{ig}, \eta_g)$ is modeled as a general nonparametric function. Thus, we also allow for arbitrary nonlinear effects of the group-level unobservable covariates. In addition, note that we do not need to specify the dimensionality of η_g, which may be either a scalar or a vector.

In many applications, it is likely that observable covariates z_{ig} and x_g are related to the unobservable η_g. Therefore, we assume that z_{ig} and x_g are endogenous in the sense that η_g is not independent of z_{ig} and x_g. Under the endogeneity of z_{ig} and x_g, $\beta(u)$ is not identified without further assumptions. To achieve identification, we assume that there exists a d_w-vector of instruments w_g satisfying $E[\varepsilon(u, \eta_g)|w_g] = 0$ (mean-independence) and $E[w_g x_g']$ is nonsingular (relevance). These are well-known conditions from the classical instrumental variable estimation theory. However, as in Hausman and Taylor (1981), since we consider the group (panel) structure of the data, some components of w_g can be *internal* instruments, that is, if some component k of z_{ig}, denoted $z_{ig}^{(k)}$, is exogenous in the sense that $E[\varepsilon(u,\eta_g)|z_{ig}^{(k)}]=0$, then we can use, say, $\sum_{1\le i\le N_g} z_{ig}^{(k)}/\sqrt{N_g}$ as an instrument. Note that one sufficient condition for mean-independence is that η_g is independent of w_g. Our mean-independence, however, is a slightly weaker condition since it is required only for a particular value u of interest.

Our problem in this paper is different from that studied in Koenker (2004), Kato et al. (2011), Galvao and Wang (2013), and Kato and Galvao (2011). Specifically, they considered a quantile panel data model which, in our framework, would be written as

$$y_{ig} = z_{ig}'\gamma(u_{ig}) + \alpha_g(u_{ig}) \tag{3.2}$$

where $\alpha_g(\cdot)$ represents the group-level fixed effect, and the researcher is interested in estimating $\gamma(\cdot)$. In a rigorous study, Kato et al. (2011) showed that their fixed effect quantile regression estimator $\hat{\gamma}(\cdot)$ is asymptotically (mean-zero) normal if the following condition holds: $G^2/\min_{g=1,\ldots,G} N_g \to 0$ as $G\to\infty$ (up-to logs). Note that this condition is stronger than that required in this paper for asymptotic (mean-zero) normality of our estimator $\hat{\beta}(\cdot)$ of $\beta(u)$ in model (3.1), which is $G^{1/2}/\min_{g=1,\ldots,G} N_g \to 0$ (up-to logs). This difference is likely to be explained by the fact that we are interested in estimating a different object. Further, Kato and Galvao (2011) suggested a fixed effect smoothed quantile regression estimator $\hat{\gamma}(\cdot)$ of $\gamma(\cdot)$ in model (3.2). They showed that their estimator is asymptotically (mean-zero) normal if $G/\min_{g=1,\ldots,G} N_g \to 0$. In addition, they derived the asymptotic bias when $\min_{g=1,\ldots,G} N_g$ is of the same order as G and obtained a bias-corrected estimator.

Our model is also different from that studied in , who considered an extension of Hausman and Taylor (1981) to cover non-linear panel data models. Formally, they considered a non-linear panel data model defined by the following equation:

$$E\left[\varphi(y_{ig}, z_{ig}'\gamma + x_g'\beta + \varepsilon_g)\right] = 0$$

where $\varphi(\cdot,\cdot)$ is a vector of moment functions and ε_g is the group-level effect. As in this paper, the authors were interested in estimating the effect of group-level covariates (coefficient β) without assuming that ε_g is independent (or mean-independent) of x_g but assuming instead that there exists an instrument w_g satisfying $E[\varepsilon_g|w_g] = 0$. Importantly, however, they assumed that $\varphi(\cdot,\cdot)$ is a vector of *smooth* functions, so that their results do not apply immediately to our model. In addition, required that $\min_{g=1,\dots,G} N_g/G > c$ for some $c > 0$ uniformly over all $G \geq 2$ to prove that their estimator is asymptotically (mean-zero) normal. In contrast, as emphasized above, we only require, up-to logs, that $\min_{g=1,\dots,G} N_g/\sqrt{G} \to \infty$ as $G \to \infty$; the improvement comes from a better control of the residuals in the Bahadur representation.

A strong assumption in our model is that the distribution of unobserved individual heterogeneity u_{ig} is not allowed to vary across groups. This assumption restricts possible interpretations of u_{ig}. However, our model does allow for the distribution of observed individual-level covariates z_{ig} to vary across groups, as well as for the unobserved group shocks to differentially affect quantiles within a group through $\varepsilon(u, \eta_g)$. Our model also allows for groups to vary because of group-level covariates. Extending our results to allow for varying distributions of unobserved individual heterogeneity would require some other assumptions or richer data structures like those in Athey and Imbens (2006).

Finally, we emphasize that in contrast with classical linear panel data models, we assume that both the number of groups G and the number of individuals in each group N_g are large. In other words, we study asymptotics where N_g goes to infinity for every $g = 1, \dots, G$ as the number of groups G increases. This gives a relevant approximation in many empirical applications.

3.3 Examples of Grouped IV Quantile Regression

To help the reader envision applications of our estimator, in this section, we provide several motivating examples of settings for which our estimator may be useful. Note that each of the following examples involves estimation of a treatment effect that varies at the group level with all endogeneity concerns also existing only at the group level.[5] As described in more detail in Section 3.4, the estimator is performed in two steps:

1. If no micro-level covariates are included, calculate the desired quantile for the outcome within each group. If micro-level variables are present, perform quantile regression for each group

[5] This is in contrast to settings where the endogeneity exists at the individual level, i.e. when the individual unobserved heterogeneity is correlated with treatment. Such situations require a different approach than the one presented here, e.g. Chernozhukov and Hansen (2005), Abadie et al. (2002), or other approaches discussed in Section 3.1.

and save the intercept term (the regression constant) from the regression for each group (analogous to partialling out in a linear model).

2. Estimate 2SLS (or OLS if the treatment is exogenous) of the group-level quantiles (or group-level quantile regression constants) on the group-level treatment and any group-level controls.

Example 1: Peer Effects of School Integration

Angrist and Lang (2004) studied how suburban student test scores were affected by the reassignment of participating urban students to suburban schools through Boston's Metco program. Before estimating their main instrumental variables model, the authors tested for a relationship between the presence of urban students in the classroom and the second decile of student test scores by estimating

$$Q_{y_{igjt}|m_{gjt},s_{gjt},\xi_{gjt}}(0.2) = \alpha_g + \beta_j + \gamma_t + \delta m_{gjt} + \lambda s_{gjt} + \xi_{gjt} \tag{3.3}$$

where $Q_{y_{igjt}|x}(0.2)$ represents the second decile of student test scores within a group, where each group is a grade $g \times$ school $j \times$ year t cell. The variables s_{gjt} and m_{gjt} denote the class size and the fraction of Metco students within each $g \times j \times t$ cell, and α_g, β_j, and γ_t represent grade, school, and year effects. The unobserved component, ξ_{gjt}, is analogous to the $\varepsilon(.2, \eta_g)$ of our model.

Angrist and Lang (2004) estimated equation (3.3) by OLS, which is equivalent to the non-IV application of our estimator with no micro-level covariates. Similar to their OLS results on average test scores, they found that classrooms with higher proportions of urban students have lower second decile test scores. Once they instrumented for a classroom's level of Metco exposure, the authors found no effect on *average* test scores. However, by not estimating model (3.3) by 2SLS, they were unable to address the causal *distributional* effects of Metco exposure.

Example 2: Occupational Licensing and Quality

Larsen (2012b) applied the estimator developed in this paper to study the effects of occupational licensing laws on the distribution of quality within the teaching profession. This application uses a difference-in-differences approach. Similar to Example 1, the explanatory variable of interest is treated as exogenous and the researcher is concerned that there may be unobserved group-level disturbances. In this application, a group is a state-year combination (s, t), and micro-level data consists of teachers within a particular state in a given year.

Let q_{ist} represent the quality of teacher i in state s who began teaching in survey year t, where quality is proxied for by a continuous measure of the selectivity of the teacher's undergraduate institution. q_{ist} is modeled as

$$q_{ist} = \gamma_s(u_{ist}) + \lambda_t(u_{ist}) + Law'_{st}\delta(u_{ist}) + \varepsilon(u_{ist}, \eta_{st}) \tag{3.4}$$

where γ_s is a state effect, λ_t is a year effect, and Law_{st} is a three-element vector containing dummies equal to one if a subject test, basic skills test, or professional knowledge test was required in state s in year t.

Because no micro-level covariates are included, the grouped regression can be obtained by simply selecting the uth quantile of quality in a given state-year cell,

$$Q_{q_{ist}|Law_{st}, \eta_{st}}(u) = \gamma_s(u) + \lambda_t(u) + Law'_{st}\delta(u) + \varepsilon(u, \eta_{st}) \tag{3.5}$$

The grouped quantile estimator is obtained via OLS on equation (3.5). Larsen (2012b) found that, for first-year teachers, occupational licensing laws requiring teachers to pass a subject test lead to a small but significant decrease in the upper tail of quality, suggestive that these laws may drive some highly qualified candidates from the occupation.

This example highlights another useful feature of grouped IV quantile regression. Including state-level effects in a standard quantile regression drastically increases the computational time, and in some cases, standard estimation packages fail to reach convergence with large numbers of nuisance parameters. However, because in this example each group (a state-year combination) corresponds to only one of these nuisance fixed effects (a state), there is no need to estimate state effects in the first-stage regression, and they can instead be estimated in the second-stage linear model.

Example 3: Distributional Effects of Suburbanization

Palmer (2011) applied the grouped quantile estimator to study the effects of suburbanization on resident outcomes. This application illustrates the use of our estimator in an IV setting. In this application, a group is a metropolitan statistical area (MSA), and individuals are MSA residents.

As an identification strategy, Palmer (2011) used the results of Baum-Snow (2007) in instrumenting suburbanization with planned highways.[6]

The model is

$$\begin{aligned}
\Delta Q_{y_{igt}|x_g,s_g,\eta_g}(u) &= \beta(u) \cdot \text{suburbanization}_g + x_g'\gamma_1(u) + \varepsilon(u, \eta_g) \\
\text{suburbanization}_g &= \pi(u) \cdot \text{planned highway rays}_g + x_g'\gamma_2(u) + v_g(u)
\end{aligned}$$

where $\Delta Q_{y|x_g,s_g,\eta_g}(u)$ is the change in the uth quantile of log wages y_{igt} within an MSA between 1950 and 1990 and x_g is a vector of controls (including a constant) conditional upon which planned highway rays$_g$ is uncorrelated with η_g and $v_g(u)$. The variable suburbanization$_g$ is a proxy measure of population decentralization, such as the amount of decline of central city population density. $\beta(u)$ is the coefficient of interest, capturing the effect of suburbanization on the within-MSA conditional wage distribution. For example, if the process of suburbanization had particularly acute effects on the prospects of low-wage workers, we may expect $\beta(u)$ to be negative for $u = 0.1$. For a given u, the grouped IV quantile approach estimates $\beta(u)$ through a 2SLS regression.

Example 4: Testing the Symmetric, Independent, Private Values Assumption in English Auctions

This example describes a test of the symmetric, independent, private values (IPV) assumption in English auction models. The example illustrates the use of the grouped quantile estimator in an IV or non-IV setting, where individual-level covariates are included. In this example, a group is a particular value of B, the number of bidders present at the auction, and an individual observation is the transaction price from a particular auction sale.[7] Assuming bidding follows the standard button auction model and bidders play the weakly dominant truth-telling equilibrium, the transaction price will be the second-order statistic of buyer valuations, and a necessary condition for the symmetric IPV setting is that quantiles of the distribution of these transaction prices should increase in the number of bidders.

To test this assumption, one can group auction sales together based on the number of bidders

[6] Baum-Snow (2007) instrumented for actual constructed highways with planned highways and estimated that each highway ray emanating out of a city caused an 18% decline in central-city population.

[7] Note that the test suggested here would not be useful if the researcher is concerned with unobservables at the individual auction level rather than at the level of the number of participants. In the former case, alternative instrumental variables quantile regression methods would be appropriate.

present at the sale (B). Let $g = 1, ..., G$ denote these groups. Let p_{ig} represent the transaction price from auction i in group g, and let B_g represent the number of bidders present at the sale. Let p_{ig} be written

$$p_{ig} = \alpha(u_{ig}) + \beta(u_{ig})B_g + \gamma(u_{ig})X_{ig} + \varepsilon(u_{ig}, \eta_g)$$

where X_{ig} represents observable auction heterogeneity. The term η_g captures group-level unobservables which specifically affect prices in auctions in which there are B_g bidders present. If the number of bidders, B_g, is treated as exogenous, grouped estimation can be performed by first doing u-quantile regression of p_{ig} on X_{ig} and a constant, separately for each group, and then regressing this constant on B_g using a linear regression. If X_{ig} is not included in the equation, our grouped estimator consists of simply selecting the uth quantile of bids for each group and regressing this quantile on B_g.[8]

If instead the number of bidders is treated as endogenous, that is, η_g affects prices as well as the number of bidders participating, then the second step of the grouped estimator would be replaced with 2SLS using an instrument for the number of bidders. For example, Haile, Hong and Shum (2003) studied US timber auctions and developed a different test of the IPV assumption which also relies on instrumenting for the number of bidders present. The authors used the numbers of nearby sawmills and logging firms as instruments for the number of bidders.

3.4 Estimator

In this section, we develop a two-stage estimator. Our main emphasis is to derive a computationally simple, yet consistent, estimator. In the first stage, for each group g and each quantile index u from the set $\mathcal{U}$ of indices of interest, we estimate u-quantile regression of y_{ig} on z_{ig} and the constant using the data $\{(y_{ig}, z_{ig}) : i = 1, ..., N_g\}$ by the classical quantile regression estimator, which can be written as

$$(\hat{\gamma}_g(u), \hat{\alpha}_g(u)) = \arg \min_{(h,a)\in\mathbb{R}^{d_z+1}} \sum_{i=1}^{N_g} \rho_u(y_{ig} - z_{ig}h - a),$$

[8] Larsen (2012a) applied this approach to data from wholesale used-car auctions and finds a positive and significant coefficient on the number of bidders present for most quantiles tested, consistent with the symmetric IPV assumption.

where $\rho_u(\cdot)$ is a "check" function of Koenker and Bassett Jr (1978), i.e. $\rho_u(x) = (u - I\{x < 0\})x$.[9] Define

$$\alpha_g(u) = x_g\beta(u) + \varepsilon(u, \eta_g),$$
$$\hat{\alpha}_g(u) = x_g\beta(u) + \varepsilon(u, \eta_g) + \nu_g(u)$$

where $\nu_g(u) = \hat{\alpha}_g(u) - \alpha_g(u)$.

The second stage consists of estimating a 2SLS regression of $\hat{\alpha}_g(u)$ on x_g using w_g as an instrument to get an estimator $\hat{\beta}(u)$ of $\beta(u)$, i.e.

$$\hat{\beta}(u) = \left(X'P_W X\right)^{-1}\left(X'P_W\hat{A}(u)\right)$$

where we use matrix notation $X = (x_1, ..., x_G)'$, $W = (w_1, ..., w_G)'$, $\hat{A}(u) = (\hat{\alpha}_1(u), ..., \hat{\alpha}_G(u))'$, and $P_W = W(W'W)^{-1}W'$. Intuitively, as the number of observations per group increases, ν_g shrinks in absolute value, and we obtain a classical instrumental variables problem. The theory presented below provides a lower bound on the growth of the number of observations per group that is sufficient to achieve consistency and asymptotic normality of $\hat{\beta}(u)$.

Note that in the first stage we run the quantile regression independently for every group. Another approach would be to estimate all parameters $\{\alpha_g(u) : g = 1, ..., G\}$ jointly. This would provide an efficiency gain given that, in this model, individual-level effects $\gamma(u)$ are group-independent. This is the approach taken up in Koenker (2004) and Kato et al. (2011) for quantile regression with panel data. Although our method is less efficient, it is computationally much less demanding since only few parameters are estimated in each regression, which can greatly reduce computation times in large datasets with many fixed effects. In addition, our approach is flexible in the sense that it yields consistent estimators even if $\gamma(\cdot)$ varies across groups.

Two special cases of the estimator are worth noting. First, if no individual-level covariates z_{ig} are included, the first stage simplifies to selecting the uth quantile of outcome variable y_{ig} within each group. Second, if x_g is independent of η_g, OLS of $\hat{\alpha}_g(u)$ on x_g may be used rather than 2SLS. When there are both no individual-level covariates and no endogeneity, the grouped IV quantile regression estimator simplifies to the minimum distance estimator described in Chamberlain (1994).

[9] Note that there may be some efficiency gains from estimating parameters jointly for the whole index set $\mathcal{U}$. For simplicity, however, we do not consider this possibility in the paper.

In summary, the steps of our estimator are

1. 1. Run u-quantile regression of y_{ig} on z_{ig} and a constant for each group, $g = 1, ..., G$. If z_{ig} is not included, simply select the uth quantile of y_{ig} within each group.

2. 2. Perform 2SLS regression, regressing the estimated constants from each group (or, with no individual-level covariates, the uth quantile of y_{ig} within each group) on x_g, instrumenting with w_g. If endogeneity is not a concern, instead perform OLS of the estimated constants on x_g.

Standard errors may then be obtained using standard approaches for robust standard errors for 2SLS or OLS (heteroskedasticity or autocorrelation-robust, clustered, etc.) as if there were no first stage, as demonstrated in Section 3.5. Section 3.5 also derives subsampling procedures that are suitable for constructing uniform confidence bands for the case when the researcher is interested in the set $\mathcal{U}$ of quantile indices u.

To conclude this section, we provide an example demonstrating that the classical quantile regression estimator is not consistent in our setting even if group-level covariates are exogenous. Consider the following model:

$$y_{ig} = \Phi^{-1}(u_{ig}) + x_g\Phi^{-1}(u_{ig}) + \eta_g \tag{3.6}$$

where observations are independent across g, $\Phi(\cdot)$ is the cdf of a $N(0, 1)$ random variable, x_g equals 1 with probability $1/2$ and equals 0 otherwise, η_g is a $N(0, 1)$ random variable that is independent of x_g, and each u_{ig} is distributed uniformly on $[0, 1]$ and is independent of η_g and across i. In this model, conditional on $x_g = 0$, y_{ig} is distributed as a $N(0, 2)$ random variable, and conditional on $x_g = 1$, y_{ig} is distributed as a $N(0, 5)$ random variable. Therefore, the uth quantile of y_{ig} given x_g equals $\sqrt{2}\Phi^{-1}(u)$ if $x_g = 0$ and equals $\sqrt{5}\Phi^{-1}(u)$ if $x_g = 0$. Therefore, running a u-quantile regression of y_{ig} on x_g (and a constant) would give an estimator of the coefficient of x_g converging in probability to $(\sqrt{5} - \sqrt{2})\Phi^{-1}(u)$, whereas the true coefficient on x_g in model (3.6) is $\Phi^{-1}(u)$. We conclude that the classical quantile regression estimator is not consistent in our setting.

3.5 Asymptotic Theory

In this section, we first formulate our assumptions. Then we present the main theoretical results of the paper.

3.5.1 Assumptions

Let $F_{ig}(\cdot)$ and $f_{ig}(\cdot)$ denote the conditional cumulative distribution function of y_{ig} and its derivative given group-level information, respectively. Let $E_g[\cdot] = E[\cdot|x_g, \eta_g]$. Denote

$$J_g(u) = E_g[(1, z_{ig}')'(1, z_{ig}')f_{ig}(\alpha_g(u) + z_{ig}'\gamma(u)|].$$

Let $c_1, C_1, c_2, C_2, c_3, C_3$ be strictly positive constants. We will assume the following regularity conditions:

A1 (Design). *Observations are independent across groups. In addition, $\{(y_{ig}, z_{ig}), i = 1, ..., N_g\}$ are independent and identically distributed group-level information.*

A2 (Covariates). $\|z_{ig}\| \leq C_1$ *and* $\|x_g\| \leq C_1$.

A3 (Instruments). *(i) For all $u \in \mathcal{U}$, $E[\varepsilon(u, \eta_g)|w_g] = 0$. (ii) $E[\|w_g\|^2] \leq C_1$. (iii) All eigenvalues of $E[w_g w_g']$ and all singular values of $E[w_g x_g']$ are bounded in absolute value away from zero by c_1.*

A4 (Coefficients). *For all $u_1, u_2 \in \mathcal{U}$, $\|\gamma(u_1) - \gamma(u_2)\| \leq C_2|u_1 - u_2|$ and $\|\beta(u_1) - \beta(u_2)\| \leq C_2|u_1 - u_2|$.*

A5 (Noise). *(i) For all $u_1, u_2 \in \mathcal{U}$, $|\varepsilon(u_1, \eta_g) - \varepsilon(u_2, \eta_g)| \leq C_2|u_1 - u_2|$. (ii) For some $\bar{u} \in \mathcal{U}$, $E[|\varepsilon(\bar{u}, \eta_g)|] \leq C_1$. (iii) $E[\|\varepsilon(\bar{u}, \eta_g)^2 w_g w_g'\|] \leq C_1$.*

A6 (Matrix $J_g(u)$). *(i) For all $u \in \mathcal{U}$, all eigenvalues of $J_g(u)$ are bounded from below by c_1. (ii) For all $u_1, u_2 \in \mathcal{U}$, $\|J_g^{-1}(u_1) - J_g^{-1}(u_2)\| \leq C_2|u_1 - u_2|^{1/2}$.*

A7 (Density). *(i) For all $u \in \mathcal{U}$, $f_{ig}(\cdot)$ is continuously differentiable in $B(\alpha_g(u) + z_{ig}'\gamma(u), c_2)$ with the derivative $f_{ig}'(\cdot)$ satisfying $|f_{ig}'(x)| \leq C_3$ for all $x \in B(\alpha_g(u) + z_{ig}'\gamma(u), c_2)$ and $|f_{ig}'(\alpha_g(u) + z_{ig}'\gamma(u)|) > c_3$. (ii) For all $u \in \mathcal{U}$ and $x \in B(\alpha_g(u) + z_{ig}'\gamma(u), c_2)$, $f_{ig}(x) \leq C_3$.*

A8 (Growth Condition). $\sqrt{G}\max_{g=1,\ldots,G}(\log(N_g)/N_g) \to 0$ *as* $G \to \infty$.

A9 (Quantile Indices). *$\mathcal{U}$ is a compact set included in $(0, 1)$.*

Assumption 1 imposes the restriction that group observations form a random sample from some population of groups. In addition, this assumption excludes conditional inter-dependence across individuals within groups given group-level information. The latter condition can be relaxed but would require more technically involved arguments. Assumption 2 can always be satisfied by a suitable transformation of the data. Assumption 3 is an identification condition. Assumptions 4 and 5 are continuity conditions. In addition, Assumption 5 requires sufficient integrability of the noise $\varepsilon(u, \eta_g)$, which is a mild regularity condition. Assumption 6 holds if all eigenvalues of $E_g[(1, z_{ig}')'(1, z_{ig}')]$ are bounded away from zero, which is a standard condition in regression analysis, and $f_{ig}(\alpha_g(u) + z_{ig}'\gamma(u))$ is bounded away from zero as well. Assumption 7 is a mild regularity condition that is often imposed in quantile regression analysis. Assumption 8 implies that the number of observations per group grows sufficiently fast and gives a particular growth rate that suffices for our results. Note, however, that our growth condition is rather weak and only requires that the number of observations per group grows as the square root of the number of groups, up-to logarithmic terms. Finally, Assumption 9 excludes quantile indices that are too close to boundaries because parameters corresponding to quantiles close to boundaries are harder to estimate. Thus, our estimator should be suitable for situations where the researcher has micro-level data on individuals within a group, and each group is assumed to come from a large population, such as a city, state, school district, etc.

3.5.2 Results

We now present our main results. We start with deriving a sub-gaussian bound for the quantile estimator (Theorem 1). This bound plays an important role in deriving the asymptotic distribution of our estimator, which is given in Theorem 2. Further, in Theorem 3, we show how to estimate the asymptotic covariance of our estimator. Finally, we provide a subsampling method to obtain uniform over $u \in \mathcal{U}$ confidence bands for the parameter of interest $\{\beta(u), u \in \mathcal{U}\}$ in Theorem 4.

Theorem 1 (Sub-gaussian tail bound). *Let Assumptions 1-9 hold. Then for all $g = 1, ..., G$,*

$$P\left(\sup_{u\in\mathcal{U}} \|(\hat{\alpha}_g(u), \hat{\gamma}_g(u)')' - (\alpha_g(u), \gamma_g(u)')'\| > x\right) \leq Ce^{-cx^2 N_g}$$

and so

$$P\left(\sup_{u\in\mathcal{U}} |\hat{\alpha}_g(u) - \alpha_g(u)| > x\right) \leq Ce^{-cx^2 N_g} \tag{3.7}$$

for some constants $c, C > 0$ that depend only on $c_1, C_1, c_2, C_2, c_3, C_3$.

Remark 1. The bound provided in Theorem 1 is *non-asymptotic*. In principle, it is also possible to calculate the exact constants in the inequality (3.7). We do not do that because it is not needed for our results. Theorem 1 also applies to a general quantile estimator and, thus, might be of independent interest. The theorem implies that large deviations of the quantile estimator from the true value are extremely unlikely. $\square$

Theorem 2 (Main convergence result). *Let Assumptions 1-9 hold. Then*

$$\sqrt{G}(\hat{\beta}(\cdot) - \beta(\cdot)) \Rightarrow \mathbb{G}(\cdot), \text{ in } \ell^{\infty}(\mathcal{U})$$

where $\mathbb{G}(\cdot)$ *is the zero-mean Gaussian process with uniformly continuous sample paths and covariance function* $C(u_1, u_2) = SJ(u_1, u_2)S'$ *where* $S = \left(Q_{xw}Q_{ww}^{-1}Q_{xw}'\right)^{-1} Q_{xw}Q_{ww}^{-1}$ *and* $J(u_1, u_2) = E[\varepsilon(u_1, \eta_g)\varepsilon(u_2, \eta_g)w_g w_g']$.

Remark 2. (i) This is our main convergence result that establishes the asymptotic behavior of our estimator. Note that we provide the *joint* asymptotic distribution of our estimator for all $u \in \mathcal{U}$. In addition, Theorem 2 implies that for any $u \in \mathcal{U}$, $\sqrt{G}(\hat{\beta}(u) - \beta(u)) \Rightarrow N(0, V)$, where $V = SJ(u, u)S'$, which is the asymptotic distribution of the classical 2SLS estimator.

(ii) In order to establish the joint asymptotic distribution of our estimator for all $u \in \mathcal{U}$, we have to deal with G independent quantile processes $\{\hat{\alpha}_g(u) - \alpha_g(u), u \in \mathcal{U}\}$. Since $G \to \infty$, classical functional central limit theorems do not apply. Therefore, we employ a non-standard but powerful Bracketing by Gaussian Hypotheses Theorem, which is also related to majorizing measures for Gaussian processes; see Theorem 2.11.11 in Van der Vaart and Wellner (1996).

(iii) In contrast with the analysis in Hausman and Taylor (1981), our model and estimators does not necessarily yield $E[\hat{\alpha}_g - \alpha_g | w_g] = 0$ because quantile estimators have finite-sample bias whereas a similar condition holds in Hausman and Taylor (1981) since the fixed effect estimator in the classical panel data model is linear. For this reason, it would seem impossible at first glance to derive a consistent estimator of $\beta(u)$ when N_g is bounded from above uniformly over all $g = 1, ..., G$ and $G > 2$. We note, however, that quantile estimators are *asymptotically* unbiased, and so we use Bahadur representation of quantile estimators to derive weak condition on the growth of N_g relative to G, so that consistent estimation of $\beta(u)$ is indeed possible. Specifically, we prove consistency under Assumption 8, which states that, up-to logs, $\min_{g=1,...,G} N_g/\sqrt{G} \to \infty$ as $G \to \infty$, which is a mild growth condition in comparison with those typically required in the analysis of quantile panel data models; see, in particular, Section 3.2. $\square$

The result in Theorem 2 derives the asymptotic behavior of our estimator. In order to perform inference, we also need an estimator of the asymptotic covariance function. For all $u_1, u_2 \in \mathcal{U}$, let

$$\hat{\mathcal{C}}(u_1, u_2) = \hat{S}\hat{J}(u_1, u_2)\hat{S}'$$

where

$$\hat{S} = (\hat{Q}_{G,xw}\hat{Q}_{G,ww}^{-1}\hat{Q}_{G,xw}')^{-1}\hat{Q}_{G,xw}\hat{Q}_{G,ww}^{-1},$$

$$\hat{J}(u_1, u_2) = \sum_{g=1}^{G}\left((\hat{\alpha}_g(u_1) - x_g'\hat{\beta}(u_1))(\hat{\alpha}_g(u_2) - x_g'\hat{\beta}(u_2))w_g w_g'\right)/G,$$

$$\hat{Q}_{G,xw} = X'W/G, \text{ and } \hat{Q}_{G,ww} = W'W/G.$$

We show that $\hat{\mathcal{C}}(u_1, u_2)$ is consistent for $\mathcal{C}(u_1, u_2)$ uniformly over $u_1, u_2 \in \mathcal{U}$.

Theorem 3 (Estimating $\mathcal{C}(\cdot,\cdot)$). *Let Assumptions 1-9 hold. Then* $\|\hat{\mathcal{C}}(u_1, u_2) - \mathcal{C}(u_1, u_2)\| = o_p(1)$ *uniformly over* $u_1, u_2 \in \mathcal{U}$.

Remark 3. Theorems 2 and 3 can be used for hypothesis testing concerning $\beta(u)$ for a given quantile index $u \in \mathcal{U}$. In particular, we have that

$$\sqrt{G}\hat{\mathcal{C}}(u, u))^{-1/2}(\hat{\beta}(u) - \beta(u)) \Rightarrow N(0, 1). \tag{3.8}$$

Importantly for applied researchers, Theorems 2 and 3 demonstrate that heteroskedasticity-robust standard errors for our estimator can be obtained by traditional White (1980) standard errors where we proceed as if $\hat{\alpha}_g(u)$ and $x_g'\hat{\beta}$ were equal to $\alpha_g(u)$ and $x_g'(u)$, respectively, that is, as if there were no first-stage estimation error. Similarly, traditional approaches for estimating autocorrelation-robust or clustered standard errors can be obtained in the usual way. □

Finally, we show how to obtain confidence bands for $\beta(u)$ that hold uniformly over $\mathcal{U}$. Let $\beta(u)_j$ and $\hat{\beta}(u)_j$ denote jth components of $\beta(u)$ and $\hat{\beta}(u)$, respectively, that is $\beta(u) = (\beta(u)_1, ..., \beta(u)_{d_x})'$ and $\hat{\beta}(u) = (\hat{\beta}(u)_1, ..., \hat{\beta}(u)_{d_x})'$. Let $D \subset 1, ..., d_x$ denote a subset of components of β a researcher is interested in. Define

$$T = \max_{j \in D}\sup_{u \in \mathcal{U}}\sqrt{G}|(\hat{\mathcal{C}}(u, u)_{jj})^{-1/2}(\hat{\beta}(u)_j - \beta(u)_j)| \tag{3.9}$$

where $\hat{\mathcal{C}}(u, u)_{jj}$ denotes the (j, j)th component of $\hat{\mathcal{C}}(u, u)$, and let $c_{1-\alpha}$ denote the $(1 - \alpha)$ quantile of T. Then uniform confidence bands of level α for $\beta(u)_j$ can be constructed as

$$\left[\hat{\beta}(u)_j - c_{1-\alpha}\left(\frac{\hat{C}(u,u)_{jj}}{G}\right)^{1/2}, \hat{\beta}(u)_j + c_{1-\alpha}\left(\frac{\hat{C}(u,u)_{jj}}{G}\right)^{1/2}\right]. \tag{3.10}$$

These confidence bands are infeasible, however, because $c_{1-\alpha}$ is unknown. We suggest estimating $c_{1-\alpha}$ by subsampling. Assume that subsample size b satisfies $b \to \infty$ but $b/G \to 0$. For a subsample of groups $\{g_1, ..., g_b\}$, run the second step of our estimator (2SLS regression) to obtain $\hat{\beta}(u, g_1, ..., g_b)$ using the data $\{x_g, w_g, \hat{\alpha}_g\}_{g\in\{g_1,...,g_b\}}$. Then calculate $T(g_1, ..., g_b)$ using equation (3.9) with $\hat{\beta}(u, g_1, ..., g_b)_j$ and $\hat{\beta}(u)_j$ replacing $\hat{\beta}(u)_j$ and $\beta(u)_j$, respectively, and b replacing G over all (or many) subsamples of size b. Then $c_{1-\alpha}$ can be estimated as the $(1-\alpha)$ empirical quantile of the sample $\{T(g_1, ..., g_b)\}$ where $\{g_1, ..., g_b\}$ varies over all (or many) subsamples of groups of size b. Denote this estimator by $\hat{c}_{1-\alpha}$. Then a feasible version of uniform confidence bands is given by equation (3.10) with $\hat{c}_{1-\alpha}$ replacing $c_{1-\alpha}$. The validity of these confidence bands is established in the following theorem.

Theorem 4 (Validity of Uniform Confidence Bands). *Let Assumptions 1-9 hold. In addition, suppose that all eigenvalues of $E[\varepsilon(u,\eta_g)^2 w_g w_g']$ are bounded away from zero uniformly over all $u \in \mathcal{U}$. Then*

$$P\left(\begin{array}{c} \beta(u)_j \in \left[\hat{\beta}(u)_j - \hat{c}_{1-\alpha}\left(\frac{\hat{C}(u,u)_{jj}}{G}\right)^{1/2}, \hat{\beta}(u)_j + \hat{c}_{1-\alpha}\left(\frac{\hat{C}(u,u)_{jj}}{G}\right)^{1/2}\right] \\ \text{for all } u \in \mathcal{U} \text{and } j \in D \end{array}\right) \to 1-\alpha$$

Remark 4. (i) *Uniform* confidence bands are typically larger than the *point-wise* confidence bands based on the result (3.8). The reason is that uniform confidence bands are constructed so that the *whole* function $\{\beta(u), u \in \mathcal{U}\}$ is contained in the bands with approximately $1-\alpha$ probability whereas point-wise bands are constructed so that for any given $u \in \mathcal{U}$, $\beta(u)$ is contained in the bands with approximately $1-\alpha$ probability. Which confidence bands to use depends on the specific purposes of the researcher.

(ii) It is also possible to estimate $c_{1-\alpha}$ by other bootstrap procedures. For example, one can use the multiplier bootstrap. Specifically, let $\epsilon_1, ..., \epsilon_G$ be an i.i.d. sequence of $N(0,1)$ random variables that are independent of the data. Then the multiplier bootstrap statistic is

$$T^B = \max_{j \in D} \sup_{u \in \mathcal{U}} \left(\hat{\mathcal{G}}(u,u)_{jj}\right)^{-1/2} \left[\hat{S} \sum_{g=1}^{G} \left(\epsilon_g(\hat{\alpha}_g(u) - x_g'\hat{\beta}(u))w_g\right) / \sqrt{G}\right]_j$$

where $[\cdot]_j$ denotes the jth component of the vector. Then the multiplier bootstrap critical value $\hat{c}^B_{1-\alpha}$ is the conditional $(1-\alpha)$ quantile of T^B given the data. Using techniques developed in Chernozhukov, Chetverikov, and Kato (2012, 2013a, 2013b), one can show that Theorem 4 holds if $\hat{c}(1-\alpha)$ is replaced by $\hat{c}^B(1-\alpha)$. □

3.6 Simulations

In order to investigate the properties of our estimator and compare to traditional quantile regression, we generate data according to the following model:

$$y_{ig} = z_{ig}\gamma(u_{ig}) + x_g\beta(u_{ig}) + \varepsilon(u_{ig}, \eta_g) \tag{3.11}$$

The variable x_g is correlated with η_g, where

$$x_g = \pi w_g + \eta_g + \nu_g \tag{3.12}$$

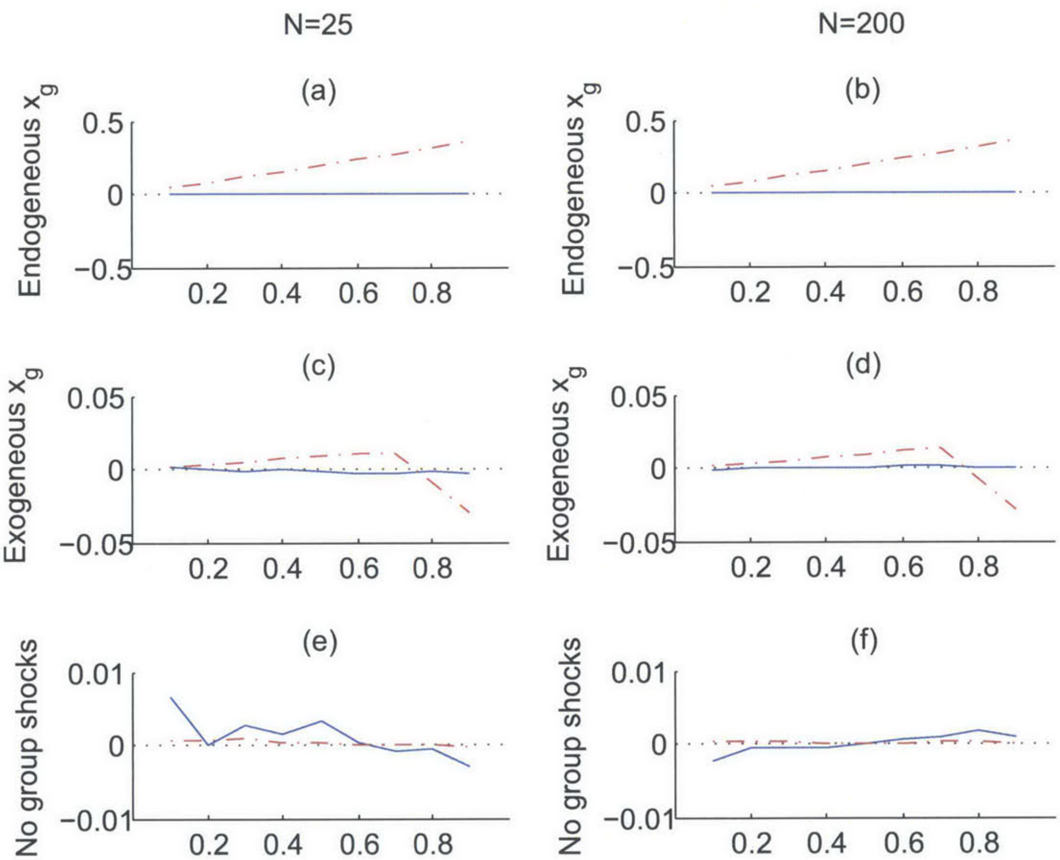
and w_g, ν_g, and z_{ig} are each distributed $\exp(0.25{*}N[0,1])$, while u_{ig} and η_g are both distributed $U[0,1]$. The quantile coefficient functions are $\gamma(u) = \beta(u) = u^{1/2}$, and $\varepsilon(u,\eta) = u\eta$. The parameter $\pi = 1$.

In addition to the case where x_g is endogenous, we also examine two special cases. First, we examine the case in which x_g is exogenous, meaning η_g does not enter (3.12). Second, we examine the case where x_g is exogenous *and* no group-level unobservables are included, meaning $\eta_g = 0$. The latter corresponds to the case in which standard quantile regression will be unbiased.

We perform our estimator and standard quantile regression for each decile ($u = 0.1, ..., 0.9$) with the number of groups (G) and the number of observations per group (N) given by (N, G) = (200,200), (200,25), (25,200), (25,25). 10,000 Monte Carlo replications were used. The results are displayed in Figures 3-1 and 3-2.

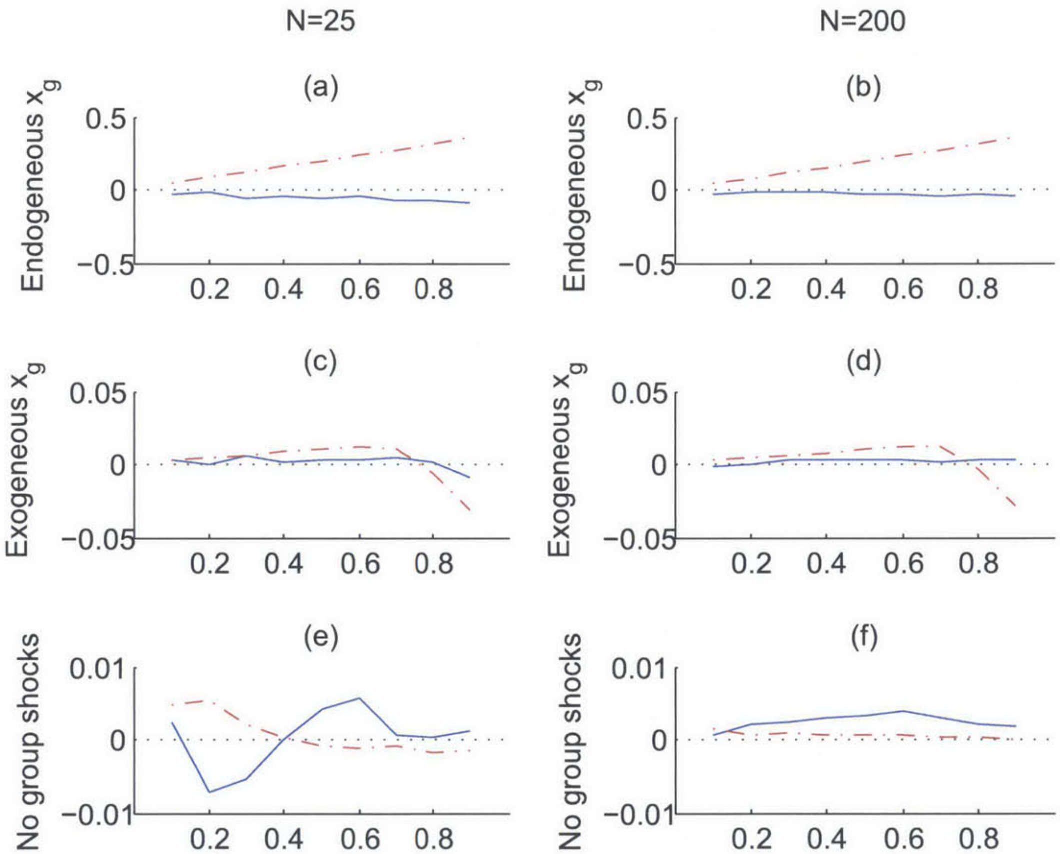
Figure 3-1 displays the simulation results with a large number of groups (G = 200). Panels on

Figure 3-1: Bias of Grouped Estimator vs. Quantile Regression; 200 Groups



Notes: Figures show mean bias for 10,000 Monte Carlo simulations with 200 groups. The number of observations per group is indicated above each column, and the group shock type is indicated by each row's y-axis labels. The solid (blue) line represents the bias of the grouped IV quantile regression estimator, the dashed (red) line represents the bias of standard quantile regression, and the dotted (black) line marks the horizontal axis. Panels (a) and (b) display the results when the treatment variable is endogenous to the group-level unobservables, panels (c) and (d) display the results when the treatment is exogenous but there still exist unobservables at the group-level, and panels (e) and (f) display the results when no group-level unobservables are present, and hence standard quantile regression should be unbiased.

Figure 3-2: Bias of Grouped Estimator vs. Quantile Regression; 25 Groups



Notes: Figures show mean bias for 10,000 Monte Carlo simulations with 25 groups. The number of observations per group is indicated above each column, and the group shock type is indicated by each row's y-axis labels. The solid (blue) line represents the bias of the grouped IV quantile regression estimator, the dashed (red) line represents the bias of standard quantile regression, and the dotted (black) line marks the horizontal axis. Panels (a) and (b) display the results when the treatment variable is endogenous to the group-level unobservables, panels (c) and (d) display the results when the treatment is exogenous but there still exist unobservables at the group-level, and panels (e) and (f) display the results when no group-level unobservables are present, and hence standard quantile regression should be unbiased.

the left of Figure 3-1 use few individuals per group ($N = 25$) and panels on the right use many individuals per group ($N = 200$). In each panel, the solid (blue) line represents the bias of the grouped IV quantile regression estimator, the dashed (red) line represents the bias of standard quantile regression, and the dotted (black) line marks the horizontal axis.

Panels (a) and (b) demonstrate that, in the presence of endogeneity, the bias of the grouped estimator is significantly lower than that of standard quantile regression. Panels (c) and (d) display the results when x_g is exogenous but η_g is still non-zero, meaning that there are still group-level unobservables (or, equivalently, left-hand-side measurement error) which standard quantile regression would not account for. In this case, standard quantile regression is biased, while the grouped estimator performs well.[10] The final panels, (e) and (f), display the results when x_g is exogenous and no group shocks are included ($\eta_g = 0$). In this situation, standard quantile regression is unbiased, and is somewhat less biased than the grouped estimator. However, the bias of the grouped estimator is also very low, as can be seen from the scale of the vertical axis. A comparison of the panels on the left to the panels on the right suggests that the bias of the grouped IV quantile estimator decreases somewhat as the number of individuals per group grows large, while a decrease in the bias of standard quantile regression is not apparent.

To illustrate the computational burden which our estimator overcomes, we re-estimated (3.11) with $\gamma(\cdot)$ and group-level fixed effects estimated jointly in one large, first-stage quantile regression (rather than estimating group-by-group quantile regression). We found that in the $(N, G) = (25, 200)$ case the joint estimation took over 35 times as long as the total time required for the group-by-group estimation. In the $(N, G) = (200, 200)$ case, the joint estimation took over 130 times as long as the group-by-group estimation.

The simulation results for the few-groups case ($G = 25$) are displayed in Figure 3-2. A similar pattern emerges to that in Figure 3-1. Even with small N and small G, the grouped estimator clearly outperforms standard quantile regression whether x_g is endogenous or exogenous, as seen in panels (a) and (c). Panel (e) demonstrates that, with no group-level shocks, quantile regression may still have bias when few individuals ($N = 25$) are included in each group. This bias decreases as N increases to 200, as seen in panel (f). Overall, the results are indicative that if group-level unobservables are present—whether they be exogenous or endogenous—the grouped estimator can have significantly less bias.

[10] For the case where $\beta(u)$ is linear in u, standard quantile regression is biased toward the median estimator, $\beta(0.5)$.

3.7 The effect of Chinese import competition on the distribution of local wages

3.7.1 Background on wage inequality

Over the past 40 years, wage inequality within the United States has increased drastically. Autor, Katz and Kearney (2008) documented that, from 1963 to 2005, the change in wages for the 90^{th} percentile earner was 55% higher than for the 10^{th} percentile earner. Economists have engaged in heated debates about the primary causes of the rising wage inequality—such as globalization, skill-biased technological change, or the declining real minimum wage—and how the importance of these factors has changed over the years. See, for example, Leamer (1994), Krugman (2000), Feenstra and Hanson (1999), Katz and Autor (1999), as well as many other papers cited in Feenstra (2010) or in Haskel, Lawrence, Leamer and Slaughter (2012).

Recent work in Autor et al. (2013) (hereafter ADH) focused on import competition and its effects on wages and employment in US local labor markets. ADH studied the period 1990–2007, when the share of US spending on Chinese imports increased dramatically from 0.6% to 4.6%. For identification, the authors used spatial variation in manufacturing concentration, showing that localized US labor markets which specialize in manufacturing were more affected by increased import competition from China. The authors found that those markets which were more exposed to increased import competition in turn had lower employment and lower wages.

We contribute to this debate by studying the effect of increased trade, in the form of increased import competition, on the distribution of local wages (rather than on the average local wages as in ADH). Given that we exploit the same variation in import competition as in ADH, we first describe the ADH framework below and then present our results.

3.7.2 Framework of Autor, Dorn, and Hanson (2013)

To study the effect of Chinese import competition on average domestic wages, ADH used Census microdata to calculate the mean wage within each Commuting Zone (CZ) in the United States.[11]

[11]The concept of commuting zones was developed by Tolbert and Sizer (1996). The United States is covered exhaustively by 722 Commuting Zones, each roughly corresponding to a local labor market.

The authors then estimated the following regression:

$$\Delta \ln w_{it} = \gamma_t + \beta_1 \Delta IPW_{it}^U + X_{it}'\beta_2 + \varepsilon_{it} \tag{3.13}$$

where $\Delta \ln w_{it}$ is the change in average log weekly wage in commuting zone i from decade t to decade $t+1$, X_{it} are characteristics of the commuting zone and decade, and γ_t are decade effects.[12] The variable of interest is ΔIPW_{it}^U, which represents the decadal change in Chinese imports per US worker, given by

$$\Delta IPW_{it}^U = \sum_j \frac{L_{ijt}}{L_{jt}} \frac{\Delta M_{jt}^{UC}}{L_{it}}$$

where L_{ijt} is the employment in period t in commuting zone i in industry j, L_{it} is the total employment in commuting zone i in period t, L_{jt} is the total US employment in industry j in period t, and ΔM_{jt}^{UC} represents the change in US imports from China in industry j from period t to $t+1$.

To address endogeneity concerns (i.e. that imports from China may be correlated with unobserved labor demand shocks), the authors instrumented for imports per worker using a measure of import exposure which replaces the change in Chinese imports to the U.S. in a given industry, ΔM_{jt}^{UC}, with the change in Chinese imports to other similarly developed nations for the same industry, ΔM_{jt}^{OC}, given by the following:[13]

$$\Delta IPW_{it}^O = \sum_j \frac{L_{ijt}}{L_{jt}} \frac{\Delta M_{jt}^{OC}}{L_{it}}.$$

Using this 2SLS approach, the authors found that a \$1,000 increase in Chinese imports per worker decreases the average log weekly wage in the commuting zone by -0.76 log points. When estimated separately by gender, the effect was more negative for males (-0.89 log points) and less so for females (-0.61 log points).[14]

[12] We have changed the notation slightly from that in ADH in order to improve clarity for our application.

[13] Specifically, ADH used Chinese imports to Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain, and Switzerland.

[14] As discussed by ADH, the presence of an extensive margin labor supply response—imports affecting not only wages but whether individuals are employed—makes these results likely a lower-bound for the effect on all workers because we don't observe wages for the unemployed population.

3.7.3 Distributional effects of increased import competition

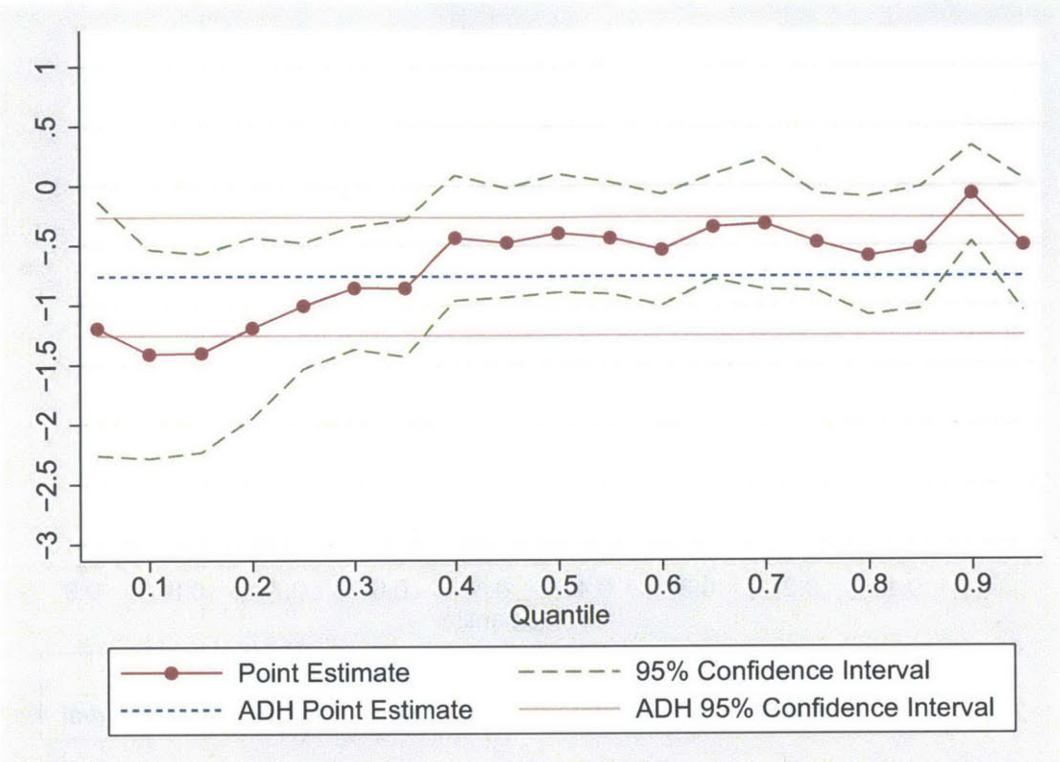
We build on the ADH framework to analyze whether low-wage earners were more adversely affected than high-wage earners by Chines import competition. To apply the grouped IV quantile regression estimator to this setting, we replace $\Delta \ln w_{it}$, the change in the average log weekly wage in the commuting zone i from period $t-1$ to t, in equation (3.13) with $\Delta \ln w_{it}^{\tau}$, the change in the τ-quantile of log wages in commuting zone i from period $t-1$ to t. Therefore, a group is this setting is a given commuting zone in a given year. We calculate these quantiles using millions of micro-level observations from the Census Integrated Public Use Micro Samples for 1990 and 2000 and the American Community Survey for 2006-2008, matching these observations to commuting zones following the strategy described in ADH. We instrument for ΔIPW_{it}^{U} using ΔIPW_{it}^{O} as described above.[15]

Several remarks are useful at this point regarding applying the grouped IV quantile framework. As explained in Section 3.1, standard quantile regression techniques would be inconsistent in this setting both because of the endogeneity of ΔIPW_{it}^{U} and because of the presence of the unobservable additive term ε_{it}. Moreover, existing methods for handling endogeneity in quantile models would not correct these issues because the endogeneity consists of a group-level treatment being correlated with the group-level unobservable additive term. Recall that existing methods for quantile regression in endogenous settings are only suited for the case where the individual-level unobserved conditional quantile itself is correlated with the treatment. Also note that, because no micro-level covariates are included in this application, the first-stage estimation simply consists of selecting the τ-quantile within each group rather than performing quantile regression for each group. Finally, note that the thought experiment behind asymptotics in this application is that the estimator is consistent as the number of groups (722 commuting zones in two decades) and the number of individuals within each group both grow large.

Figure 3-3, 3-4, and 3-5 display the results of the grouped IV quantile regression estimator for the full sample, for males only, and for females only. Each figure displays $\tau-$quantile estimates for $\tau \in \{0.5, 1, ..., 0.95\}$, along with pointwise 95% confidence bands about each estimate. The figures also display the 2SLS effect found in ADH and 95% confidence intervals corresponding to their IV estimate of Chinese import penetration on the change in CZ-level average wages.

[15]We also follow ADH by weighting observations by CZ population in period $t-1$ and by clustering at the state level.

Figure 3-3: Effect of Chinese Import Competition on Conditional Wage Distribution: Full Sample

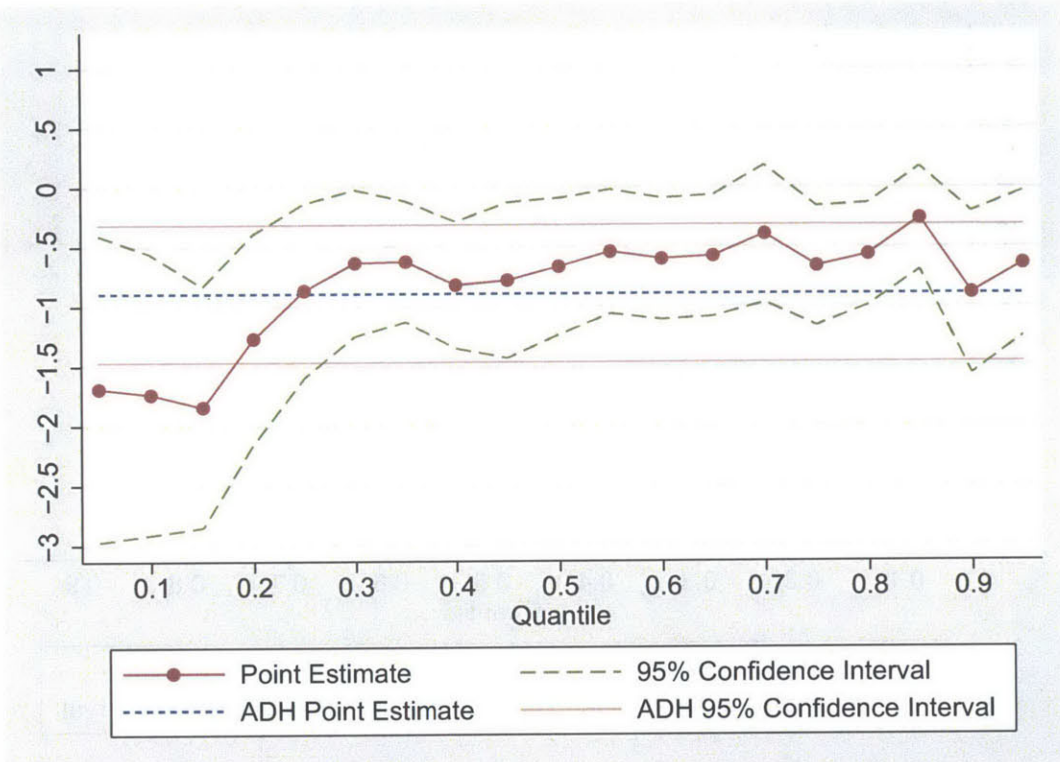


Notes: Figure plots grouped IV quantile regression estimates of the effect of a $1,000 increase in Chinese imports per worker on the conditional wage distribution (β_1 in equation (3.13) in the text when the change in average log wages for Commuting Zone i between time period $t-1$ and t $\Delta \ln w_{it}$ is replaced with the change in the τ-quantile of log wages $\Delta \ln w_{it}^{\tau}$). The dashed horizontal line is the ADH estimate of β_1 in equation (3.13). 95% pointwise confidence intervals are constructed from robust standard errors clustered by state and observations are weighted by CZ population, as in ADH. Units on the vertical axis are log points.

Each figure provides evidence that Chinese import competition affected the wages of low-wage earners more than high-wage earners. For all three samples, the estimated causal effect of Chinese import penetration is less negative for higher quantiles of the conditional wage distribution. Furthermore, in each case, we can reject an effect size of zero for all quantiles below the median but cannot for all quantiles above the median. The point estimates are suggestive that the average negative effect of Chinese import penetration is primarily driven by large negative effects for those in the lower tail, where the effect is twice as large in magnitude as the average effect. For example, Figure 3-3 demonstrates that for most wage-earners (from the 0.35 quantile and above) the effect of Chinese import competition was one-third smaller in magnitude than the effect on the average estimated by ADH.

Despite this distributional heterogeneity, for the full sample of wage-earners (Figure 3-3) and for

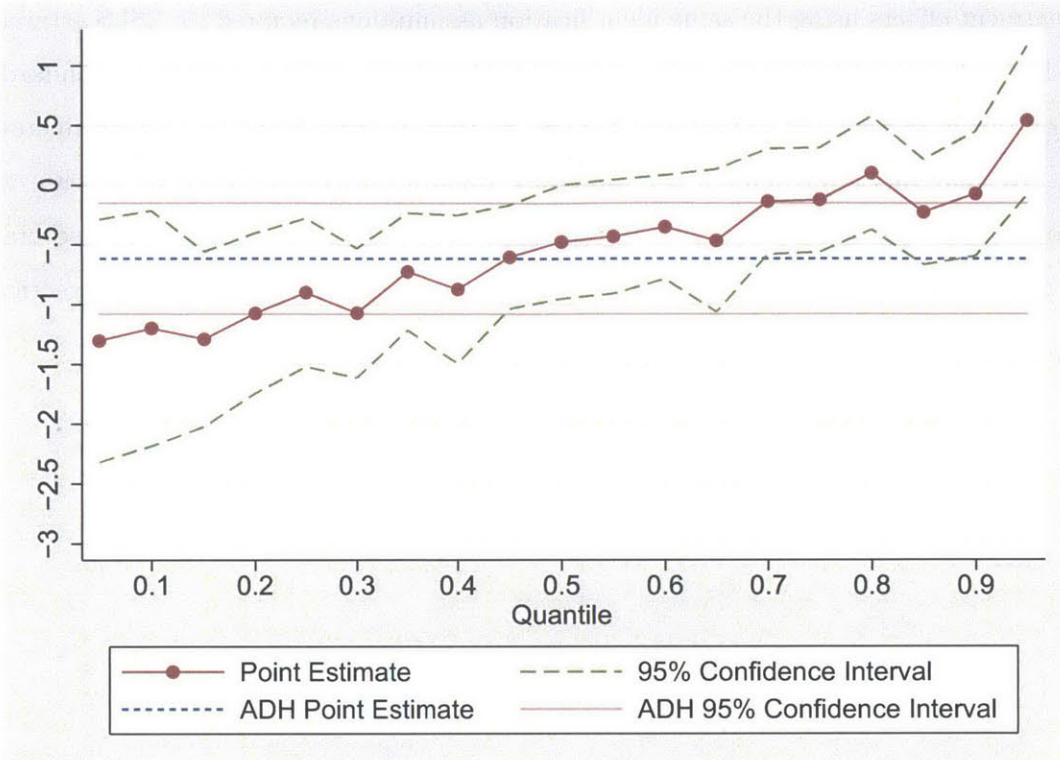
Figure 3-4: Effect of Chinese Import Competition on Conditional Wage Distribution: Males Only



Notes: Figure plots grouped IV quantile regression estimates for the male-only sample of the effect of a \$1,000 increase in Chinese imports per worker on the male conditional wage distribution (β_1 in equation (3.13) in the text when the change in average log wages for Commuting Zone i between time period $t-1$ and t $\Delta \ln w_{it}$ is replaced with the change in the τ-quantile of log wages $\Delta \ln w_{it}^{\tau}$). The dashed horizontal line is the ADH estimate of β_1 in equation (3.13). 95% pointwise confidence intervals are constructed from robust standard errors clustered by state and observations are weighted by CZ population, as in ADH. Units on the vertical axis are log points.

the male-only sample (Figure 3-4), the effect on the average estimated by ADH is contained in the confidence band of the quantile estimates at nearly every quantile, implying that it would be difficult to reject that there is indeed a constant, negative effect of Chinese import competition on wages at each point in the wage distribution. By contrast, in the female-only sample (Figure 3-5), most quantiles above 0.7 yield positive point estimates, with a 95% confidence band lying strictly above the estimated average effect from ADH. This is suggestive not only that low-earning females were more negatively affected than high-earning females by increased import competition, but that wages at the top of the female conditional wage distribution may even have increased in response to import penetration. This example highlights a distinction which would have been missed if one were to focus only on the average impact of import competition.

Figure 3-5: Effect of Chinese Import Competition on Conditional Wage Distribution: Females Only



Notes: Figure plots grouped IV quantile regression estimates for the female-only sample of the effect of a \$1,000 increase in Chinese imports per worker on the female conditional wage distribution (β_1 in equation (3.13) in the text when the change in average log wages for Commuting Zone i between time period $t-1$ and t $\Delta \ln w_{it}$ is replaced with the change in the τ-quantile of log wages $\Delta \ln w_{it}^{\tau}$). The dashed horizontal line is the ADH estimate of β_1 in equation (3.13). 95% pointwise confidence intervals are constructed from robust standard errors clustered by state and observations are weighted by CZ population, as in ADH. Units on the vertical axis are log points.

3.8 Conclusion

In this paper, we present a quantile extension of Hausman and Taylor (1981), providing an approach for estimating distributional effects of a group-level treatment variable. The estimator is computationally simple, consisting of the following steps: 1) for each group, perform quantile regression of the outcome on individual-level covariates and on a constant; 2) perform a linear regression of the group-level constants on the group-level treatment. If the treatment is correlated with group-level unobservables, this linear regression should be a 2SLS regression. We provide a rigorous derivation of the asymptotic properties of this estimator and demonstrate its consistency and asymptotic normality, as well as provide an approach for obtaining uniform confidence bands.

An empirical application to the setting of Autor et al. (2013) highlights the usefulness of our estimator in providing a method to detect the presence of distributional heterogeneity in group-level treatment effects using the same identification assumptions required for 2SLS estimates to be valid. Monte Carlo simulations show that this estimator has much lower bias than standard quantile regression in the presence of the group-level shocks that are ubiquitous in applied microeconomic models. We find this improvement over standard quantile regression both in the case where the group-level treatment is exogenous and in the case where it is endogenous. Finally, several applied microeconomics examples illustrate the potential of this estimator for widespread practical use.

3.A Proofs Appendix

In this Appendix, we first prove some preliminary lemmas. Then we present the proofs of the theorems stated in the main part of the text. In all proofs, c and C denote strictly positive generic constants that depend only on $c_1, C_1, c_2, C_2, c_3, C_3$ but whose value can change at each appearance.

We will use the following additional notation. Denote $Z_{ig} = (1, z_{ig}')'$, $\delta_g(u) = (\alpha_g(u), \gamma(u)')'$, and $\hat{\delta}_g(u) = (\hat{\alpha}_g(u), \hat{\gamma}(u)')'$. Let $A(u) = (\alpha_1(u), ..., \alpha_G(u))'$ and

$$\tilde{\beta}(u) = (X'P_W X)^{-1}(X'P_W A(u)). \tag{14}$$

For $\eta, \delta \in \mathbb{R}^{d_z+1}$, and $u \in \mathcal{U}$, consider the function $f_{\eta,\delta,u} : \mathbb{R}^{d_z} \times \mathbb{R} \to \mathbb{R}$ defined by

$$f_{\eta,\delta,u}(z, y) = (1, z')\eta(1\{y \leq (1, z')\delta\} - u).$$

Let $\mathcal{F} = \{f_{\eta,\delta,u} : \eta, \delta \in \mathbb{R}^{d_z+1}; u \in \mathcal{U}\}$, i.e. $\mathcal{F}$ is the class of functions $f_{\eta,\delta,u}$ as η, δ vary over $\mathbb{R}^{d_z+1}$ and u varies over $\mathcal{U}$. For $\delta \in \mathbb{R}^{d_z+1}$ and $u \in \mathcal{U}$, let the function $h_{\delta,u} : \mathbb{R}^{d_z} \times \mathbb{R} \to \mathbb{R}^{d_z+1}$ be defined by

$$h_{\delta,u}(z, y) = (1, z')'(1\{y \leq (1, z')\delta\} - u).$$

Let $h_{k,\delta,u}$ denote kth component of $h_{\delta,u}$. Let $\mathcal{H}_k = \{h_{k,\delta,u} : \delta \in \mathbb{R}^{d_z+1}; u \in \mathcal{U}\}$. Note that $\mathcal{H}_k \subset \mathcal{F}$ for all $k = 1, ..., d_z + 1$.

We will also use the following notation from the empirical process literature,

$$\mathbb{G}^g(f) = \sum_{i=1}^{N_g} \left(f(z_{ig}, y_{ig}) - E_g[f(z_{ig}, y_{ig})]\right) / \sqrt{N_g}$$

for $f \in \mathcal{F}$ or $f \in \mathcal{H}_k$.

Preliminary Lemmas

In all lemmas, we implicitly impose Assumptions 1-9.

Lemma 1. *There exist constants $c, C > 0$ such that*

$$\|E_g[h_{\delta,u}(z_{ig}, y_{ig})] - J_g(u)(\delta - \delta_g(u))\| \leq C\|\delta - \delta_g(u)\|^2, \tag{15}$$

$$E_g[(\delta - \delta_g(u))' h_{\delta,u}(z_{ig}, y_{ig})] \geq c\|\delta - \delta_g(u)\|^2. \tag{16}$$

for all $u \in \mathcal{U}$ and $\delta \in \mathbb{R}^{d_z+1}$ satisfying $\|\delta - \delta_g(u)\| \leq c$.

Proof. Second-order Taylor expansion around $\delta_g(u)$ gives

$$\begin{aligned} E_g[h_{\delta,u}(z_{ig}, y_{ig})] &= E_g[Z_{ig}(1\{y_{ig} \leq Z_{ig}'\delta\} - u)] = E_g[Z_{ig}(F_g(Z_{ig}'\delta) - u)] \\ &= E_g[Z_{ig}(F_g(Z_{ig}'\delta_g(u)) - u)] + J_g(u)(\delta - \delta_g(u)) + r_n(u), \end{aligned}$$

where $r_n(u)$ is the remainder. The first claim of the lemma follows from the facts that $E_g[Z_{ig}(F_g(Z_{ig}'\delta_g(u)) - u)] = 0$, which holds because $Z_{ig}'\delta_g(u)$ is the conditional uth quantile of y_{ig}, and that $\|r_n(u)\| \leq C\|\delta - \delta_g(u)\|^2$ for some $C > 0$ by Assumptions 2 and 7.

To prove the second claim, note that if $\|\delta - \delta_g(u)\|$ is sufficiently small, then $\|(\delta - \delta_g(u))' r_n(u)\| \leq c\|\delta - \delta_g(u)\|^2$ for an arbitrarily small constant $c > 0$. On the other hand,

$$(\delta - \delta_g(u))' J_g(u)(\delta - \delta_g(u)) \geq c\|\delta - \delta_g(u)\|^2$$

for some $c > 0$ by Assumption 6. Combining these inequalities gives the second claim. □

Lemma 2. *Consider the function class $\mathcal{F}$ defined in the beginning of the Appendix. We have that $\mathcal{F}$ is a VC subgraph class of functions. Moreover, for all $k = 1, ..., d_z + 1$, $\mathcal{H}_k$ is a VC subgraph class of functions as well.*

Proof. A similar proof can be found in Belloni, Chernozhukov and Hansen (2006). We present the proof here for the sake of completeness. Consider the class of sets $\{x \in \mathbb{R}^{d_z+2} : a'x \leq 0\}$ with a varying over $\mathbb{R}^{d_z+2}$. It is well known that this is a VC subgraph class of sets, see, for example, exercise 14 of chapter 2.6 in Van der Vaart and Wellner (1996). Further, note that

$$\begin{aligned} \{(z, y, t) : f_{\eta,\delta,u}(z, y) > t\} &= \left(\{y \leq (1, z')\delta\} \cap \{(1, z')\eta > t/(1-u)\}\right) \\ &\quad \cup \left(\{y > (1, z')\delta\} \cap \{(1, z')\eta < -t/u\}\right). \end{aligned}$$

Therefore, the first result follows from Lemma 2.6.17, parts ii and iii, in Van der Vaart and Wellner (1996). The second result follows from the fact that $\mathcal{H}_k \subset \mathcal{F}$. $\square$

Lemma 3. *There exist constants $c, C > 0$ such that*

$$E_g\left[\sup_{u\in\mathcal{U}:|u-v|\leq\epsilon} \|\mathbb{G}^g(h_{\delta_g(u),u}) - \mathbb{G}^g(h_{\delta_g(v),v})\|^2\right] \leq C\epsilon$$

for all $\epsilon \in (0, c)$ and $v \in \mathcal{U}$.

Proof. We have

$$E_g\left[\sup_{u\in\mathcal{U}:|u-v|\leq\epsilon} \|\mathbb{G}^g(h_{\delta_g,u}) - \mathbb{G}^g(h_{\delta_g(v),v})\|^2\right]$$
$$\leq \sum_{k=1}^{d_z+1} E_g\left[\sup_{u\in\mathcal{U}:|u-v|\leq\epsilon} |\mathbb{G}^g(h_{k,\delta_g,u}) - \mathbb{G}^g(h_{k,\delta_g(v),v})|^2\right].$$

Consider the function $F : \mathbb{R}^{d_z} \times \mathbb{R} \to \mathbb{R}$ given by

$$F(z, y) = C\left(1\{|y - (1, z')\delta_g(v), v| \leq C\epsilon\} + \epsilon\right)$$

for some sufficiently large $C > 0$. By Assumptions 2, 4, and 5, $\|\delta_g(u) - \delta_g(v)\| \leq C|u - v|$, and so $|Z'_{ig}(\delta_g(u) - \delta_g(v))| \leq C|u - v|$ for some $C > 0$. Therefore, for all $u \in \mathcal{U}$ satisfying $|u - v| \leq \epsilon$,

$$\left|h_{k,\delta_g(u),u}(z_{ig}, y_{ig}) - h_{k,\delta_g(v),v}(z_{ig}, y_{ig})\right| \leq F(z_{ig}, y_{ig}).$$

Note that $E_g[F^2(z_{ig}, y_{ig})] \leq C\epsilon$ for some $C > 0$ by Assumption 7. By Lemma 2, $\mathcal{H}_k$ is a VC subgraph class of functions. So, applying Theorem 6, part (ii), from Appendix B with F as an envelope yields

$$E_g\left[\sup_{u\in\mathcal{U}:|u-v|\leq\epsilon} |\mathbb{G}^g(h_{k,\delta_g,u}) - \mathbb{G}^g(h_{k,\delta_g(v),v})|^2\right] \leq C\epsilon$$

for all $k = 1, ..., d_z + 1$ and some $C > 0$. The claim of the lemma follows. $\square$

Lemma 4. *There exist constants $c, C > 0$ such that*

$$E_g\left[\sup_{u\in\mathcal{U}}\sup_{\delta\in\mathbb{R}^{d_z+1}:\|\delta-\delta_g(u)\|\le\epsilon}\|\mathbb{G}^g(h_{\delta,u})-\mathbb{G}^g(h_{\delta_g(u),u})\|\right]$$
$$\le C(\epsilon+1/\sqrt{N_g})\sqrt{\log\left(1/(\epsilon+1/\sqrt{N_g})\right)}$$

for all $\epsilon \in (0, c)$.

Proof. We have

$$E_g\left[\sup_{u\in\mathcal{U}}\sup_{\delta\in\mathbb{R}^{d_z+1}:\|\delta-\delta_g(u)\|\le\epsilon}\|\mathbb{G}^g(h_{\delta,u})-\mathbb{G}^g(h_{\delta_g(u),u})\|\right] \tag{17}$$
$$\le C\sum_{k=1}^{d_z+1}E_g\left[\sup_{u\in\mathcal{U}}\sup_{\delta\in\mathbb{R}^{d_z+1}:\|\delta-\delta_g(u)\|\le\epsilon}|\mathbb{G}^g(h_{k,\delta,u})-\mathbb{G}^g(h_{k,\delta_g(u),u})|\right] \tag{18}$$

for some $C > 0$. Consider the function class

$$\tilde{\mathcal{H}}_k = \{h_{k,\delta,u}-h_{k,\delta_g(u),u} : u\in\mathcal{U}; \|\delta-\delta_g(u)\|\le\epsilon\}.$$

We have that $\tilde{\mathcal{H}}_k \subset \mathcal{F}-\mathcal{F}$. Therefore, it follows from Lemma 2 that $\tilde{\mathcal{H}}_k$ has a polynomial uniform entropy. In addition, all functions from $\tilde{\mathcal{H}}_k$ are bounded in absolute value by some constant $C > 0$. Applying Theorem 6, part (i), from Appendix B yields

$$E_g\left[\sup_{f\in\tilde{\mathcal{H}}_k}|\mathbb{G}^g(f)|\right] \le CE_g\left[\chi_n+\int_0^{\chi_n\wedge 1}\sqrt{1+\log(1/x)}dx\right]$$

for some $C > 0$ where $\chi_n = \sup_{f\in\tilde{\mathcal{H}}_k}\sum_{i=1}^{N_g}|f(z_{ig},y_{ig})|/N_g$. Further, integrating by parts gives

$$\int_0^X\sqrt{\log(1/x)}dx = X\sqrt{\log(1/X)}+\frac{1}{2}\int_0^X 1/\sqrt{\log(1/x)}dx \le CX\sqrt{\log(1/X)}$$

for some $C > 0$ and all $X \in (0,1)$. Therefore,

$$E_g\left[\sup_{f\in\tilde{\mathcal{H}}_k}|\mathbb{G}^g(f)|\right] \le CE_g\left[\chi_n+(\chi_n\wedge 1)\sqrt{\log(1/(\chi_n\wedge 1))}\right] \tag{19}$$
$$\le C\left(E_g[\chi_n]+E_g[\chi_n\wedge 1]\sqrt{\log(1/E_g[\chi_n\wedge 1])}\right) \tag{20}$$

where in the last line we used the fact that the function $l(x) = x\sqrt{\log(1/x)}$ is concave. Indeed, calculations show that

$$
\begin{aligned}
l'(x) &= \sqrt{\log(1/x)} - 1/\left(2\sqrt{\log(1/x)}\right) \\
l''(x) &= -1/\left(2x\sqrt{\log(1/x)}\right) - 1/\left(4x\log^{3/2}(1/x)\right) \leq 0.
\end{aligned}
$$

Now, to bound $E_g[\chi_n]$, we apply Theorem 6, part (ii), from Appendix B. This theorem is applicable because the function class $\{|f| : f \in \tilde{\mathcal{H}}_k\}$ also has polynomial uniform entropy since for any random variables X and Y, we have

$$
\begin{aligned}
E[(|X| - |Y|)^2] &= E[X^2] + E[Y^2] - 2E[|XY|] \\
&\leq E[X^2] + E[Y^2] - 2E[XY] \leq E[(X - Y)^2].
\end{aligned}
$$

This gives

$$
E_g[\chi_n] \leq C\left(\epsilon + 1/\sqrt{N_g}\right) \tag{21}
$$

for some $C > 0$ where we also used the fact that for all $f \in \tilde{\mathcal{H}}_k$, $E_g[|f(z_{ig}, y_{ig}|] \leq C\epsilon$. Combining (21) with (17), (18), (19), and (20) gives the asserted claim. □

Lemma 5. *Uniformly over $u \in \mathcal{U}$,*

$$
-\sum_{g=1}^{G} J_g^{-1}(u)\mathbb{G}^g(h_{\delta_g(u),u})w_g'/\sqrt{G} = O_p(1).
$$

Proof. To prove this lemma, we apply Theorem 5 from Appendix B with the semi-metric $\rho(u_1, u_2) = C\sqrt{|u_1 - u_2|}$ for $u_1, u_2 \in \mathcal{U}$ and some sufficiently large constant $C > 0$. Clearly, ρ is Gaussian-dominated. Define $v_g(u) = J_g^{-1}(u)\mathbb{G}^g(h_{\delta_g(u),u})$ and

$$
Z_{k,m,g}(u) = v_{k,g}(u)w_{m,g}/\sqrt{G}
$$

where $v_{k,g}(u)$ and $w_{m.g}$ denote kth and mth components of $v_g(u)$ and w_g correspondingly. Then for any $a > 0$,

$$\sum_{g=1}^{G} E\left[\sup_{u\in\mathcal{U}} |Z_{k,m,g}(u)| 1\left\{\sup_{u\in\mathcal{U}} |Z_{k,m,g}(u)| > a\right\}\right]$$
$$\leq \sum_{g=1}^{G} E\left[\sup_{u\in\mathcal{U}} Z^2_{k,m,g}(u) 1\left\{\sup_{u\in\mathcal{U}} |Z_{k,m,g}(u)| > a\right\}\right] /a$$
$$\leq E\left[\sup_{u\in\mathcal{U}} (v_{k,g}(u) w_{m,g})^2 1\left\{\sup_{u\in\mathcal{U}} |v_{k.g}(u) w_{m.g}| > \sqrt{G}a\right\}\right] /a$$
$$\leq CE\left[\sup_{u\in\mathcal{U}} \|\mathbb{G}^g(h_{\delta_g(u).u})\|^2 w^2_{m,g} 1\left\{\sup_{u\in\mathcal{U}} \|\mathbb{G}^g(h_{\delta_g(u),u})\| w_{m,g} > \sqrt{G}a\right\}\right] /a$$
$$\leq \sum_{j,l=1}^{d_z+1} CE\left[\sup_{u\in\mathcal{U}} \mathbb{G}^g(h_{j,\delta_g(u),u})^2 w^2_{m.g} 1\left\{\sup_{u\in\mathcal{U}} \mathbb{G}^g(h_{l,\delta_g(u),u}) w_{m.g} > \sqrt{G}a/\sqrt{d_z+1}\right\}\right] /a = L$$

for some $C > 0$ where we used Assumption 6. Note that $h_{j,\delta_g(u).u} \in \mathcal{F}$ and for some $C > 0$, $|h_{j,\delta_g(u),u}(z_{ig}, y_{ig})| \leq C$ for all $j = 1, ..., d_z + 1$. Therefore, combining Lemma 2 and Theorem 6, part (ii), from Appendix B yields $E_g[\sup_{u\in\mathcal{U}} \mathbb{G}^g(h_{j,\delta_g(u),u})^2] \leq C$. Combing this inequality with Assumption 3 implies that $E[\sup_{u\in\mathcal{U}} \mathbb{G}^g(h_{j.\delta_g(u),u})^2 w^2_{m,g}] \leq C$ for some $C > 0$. So, by the Dominated Convergence Theorem, $L \to 0$ as $G \to \infty$, which gives the condition (i) of Theorem 5.

Further, for any $u_1, u_2 \in \mathcal{U}$,

$$\begin{aligned}\sum_{g=1}^{G} E\left[(Z_{k,m,g}(u_1) - Z_{k,m,g}(u_2))^2\right] &\leq E\left[(v_{k.g}(u_1) - v_{k,g}(u_2))^2 w^2_{m,g}\right] \\ &\leq CE\left[\|J_g^{-1}(u_1) - J_g^{-1}(u_2)\|^2 \|\mathbb{G}^g(h_{\delta_g(u_1),u_1})\|^2 w^2_{m.g}\right] \\ &\quad + CE\left[\|J_g^{-1}(u_2)\|^2 \|\mathbb{G}^g(h_{\delta_g(u_1).u_1} - h_{\delta_g(u_2),u_2})\|^2 w^2_{m,g}\right]\end{aligned}$$

for some $C > 0$. Since $E_g[\|\mathbb{G}^g(h_{\delta_g(u_1).u_1})\|^2] \leq C$ and $\|J_g^{-1}(u_1) - J_g^{-1}(u_2)\|^2 \leq C|u_1 - u_2|$ for some $C > 0$ by Assumption 6, we have

$$E\left[\|J_g^{-1}(u_1) - J_g^{-1}(u_2)\|^2 \|\mathbb{G}^g(h_{\delta_g(u_1),u_1})\|^2 w^2_{m,g}\right] \leq C|u_1 - u_2|$$

for some $C > 0$. Further, $\|J_g^{-1}(u_2)\| \leq C$ by Assumption 6 and

$$E_g\left[\|\mathbb{G}^g(h_{\delta_g(u_1),u_1} - h_{\delta_g(u_2),u_2})\|^2\right] \leq \sum_{k=1}^{d_z+1} E[(h_{k,\delta_g(u_1),u_1} - h_{k,\delta_g(u_2),u_2})^2] \leq C|u_1 - u_2|.$$

Therefore,

$$E\left[\|J_g^{-1}(u_2)\|^2\|\mathbb{G}^g(h_{\delta_g(u_1),u_1}-h_{\delta_g(u_2),u_2})\|^2 w_{m,g}^2\right]\leq C|u_1-u_2|$$

for some $C>0$, and so

$$\sum_{g=1}^{G}E\left[(Z_{k,m,g}(u_1)-Z_{k,m,g}(u_2))^2\right]\leq C|u_1-u_2|\leq\rho^2(u_1,u_2),$$

which is condition (ii) of Theorem 5.

Finally, condition (iii) of Theorem 5 holds because for any $\epsilon>0$ and $v\in\mathcal{U}$,

$$\begin{aligned}
&\sup_{t>0}\sum_{g=1}^{G}t^2P\left(\sup_{u\in\mathcal{U}:\rho(u,v)\leq\epsilon}|Z_{k,m,g}(u)-Z_{k,m,g}(v)|>t\right)\\
&\leq\sum_{g=1}^{G}E\left[\sup_{u\in\mathcal{U}:\rho(u,v)\leq\epsilon}|Z_{k,m,g}(u)-Z_{k,m,g}(v)|^2\right]\\
&\leq\epsilon^2/2+CE\left[\sup_{u\in\mathcal{U}:\rho(u,v)\leq\epsilon}\|\mathbb{G}^g(h_{\delta_g(u),u}-h_{\delta_g(v),v})\|^2w_{m,g}^2\right]\leq\epsilon^2
\end{aligned}$$

for some $C>0$ where we used Markov inequality in the second line, and the same argument as in the verification of the second condition and Lemma 3 in the third line (recall also that the constant in the definition of ρ is sufficiently large).

The claim of the lemma now follows by applying Theorem 5. □

Proofs of Theorems

Proof of Theorem 1. Since $\rho_u(\cdot)$ is convex, for $x>0$, $\|\hat{\delta}_g(u)-\delta_g(u)\|\leq x$ for all $u\in\mathcal{U}$ if and only if

$$\inf_{u\in\mathcal{U}}\inf_{\eta\in\mathbb{R}^{d_z+1};\|\eta\|=1}\sum_{i=1}^{N_g}f_{\eta,\delta_g(u)+x\eta,u}(z_{ig},y_{ig})/N_g\geq 0. \tag{22}$$

Now, it follows from Lemma 1 (inequality (16)) that inequality (22) holds if

$$\inf_{u\in\mathcal{U}}\inf_{\eta\in\mathbb{R}^{d_z+1};\|\eta\|=1}\sum_{i=1}^{N_g}\left(f_{\eta,\delta_g(u)+x\eta,u}(z_{ig},y_{ig})-E_g[f_{\eta,\delta_g(u)+x\eta,u}(z_{ig},y_{ig})]\right)/N_g\geq -cx,$$

for some sufficiently small $c > 0$, which in turn follows if

$$\inf_{u\in\mathcal{U}} \inf_{\eta,\delta\in\mathbb{R}^{d_z+1};\|\eta\|=1} \mathbb{G}^g(f_{\eta,\delta,u}) \geq -cx\sqrt{N_g}. \tag{23}$$

Note that for any $\eta \in \mathbb{R}^{d_z+1}$ satisfying $\|\eta\| = 1$, $|f_{\eta,\delta,u}| \leq 2\|Z_{ig}\| \leq C$ for some $C > 0$ by Assumption 2. In addition, it follows from Lemma 2 and Theorem 2.6.7 in Van der Vaart and Wellner (1996) that the conditions of Theorem 7 hold for the function class $\{f_{\eta,\delta,u} \in \mathcal{F} : u \in \mathcal{U}; \eta, \delta \in \mathbb{R}^{d_z+1}; \|\eta\| = 1\}$. Therefore, Theorem 7 shows that inequality (23) holds with probability not smaller than

$$1 - C\exp(-cx^2 N_g)$$

for some $c, C > 0$. The asserted claim follows. □

Proof of Theorem 2. The proof consists of two steps. First, we show that $\sqrt{G}(\hat{\beta}(u) - \tilde{\beta}(u)) = o_p(1)$ uniformly over $u \in \mathcal{U}$ where $\tilde{\beta}(u)$ is defined in (14). Second, we show that $\sqrt{G}(\tilde{\beta}(\cdot) - \beta(\cdot)) \Rightarrow \mathbb{G}(\cdot)$ in $\ell^\infty(\mathcal{U})$. Combining these steps gives the result.

Denote $\hat{Q}_{G,xw} = X'W/G$ and $\hat{Q}_{G,ww} = W'W/G$. Then

$$\sqrt{G}(\hat{\beta}(u) - \tilde{\beta}(u)) = \left(\hat{Q}_{G,xw}\hat{Q}_{G,ww}^{-1}\hat{Q}_{G,xw}'\right)^{-1} \hat{Q}_{G,xw}\hat{Q}_{G,ww}^{-1}\left(W'(\hat{A}(u) - A(u))/\sqrt{G}\right).$$

By Assumptions 1 and 3, $X'W/G \to_p E[x_g w_g']$, $W'W \to_p E[w_g w_g']$, and

$$\hat{S} = \left(\hat{Q}_{G,xw}\hat{Q}_{G,ww}^{-1}\hat{Q}_{G,xw}'\right)^{-1} \hat{Q}_{G,xw}\hat{Q}_{G,ww}^{-1} \to_p \left(Q_{xw}Q_{ww}Q_{xw}'\right)^{-1} Q_{xw}Q_{ww}^{-1} = S \tag{24}$$

where $Q_{xw} = E[x_g w_g']$ and $Q_{ww} = E[w_g w_g']$. Therefore, to prove the first step, it suffices to show that

$$S(u) = \sum_{g=1}^{G}(\hat{\delta}_g(u) - \delta_g(u))w_g'/\sqrt{G} = o_p(1)$$

uniformly over $u \in \mathcal{U}$ (recall that $\hat{\alpha}_g(u) - \alpha_g(u)$ is the first component of $\hat{\delta}_g(u) - \delta_g(u)$).

Write $S(u) = S_1(u) + S_2(u)$ where

$$\begin{aligned} S_1(u) &= -\sum_{g=1}^{G} J_g^{-1}(u)\mathbb{G}^g(h_{\delta_g(u),u})w_g'/\sqrt{GN_g}, \\ S_2(u) &= \sum_{g=1}^{G}\left(J_g^{-1}(u)\mathbb{G}^g(h_{\delta_g(u),u}) + \sqrt{N_g}(\hat{\delta}_g(u) - \delta_g(u))\right) w_g'/\sqrt{GN_g}. \end{aligned}$$

Consider $S_1(u)$. Combining Assumptions 2, 3, 6, and 8 gives

$$E[\|S_1(u)\|^2] \leq C/\min_{g=1,\dots,G} N_g \to 0.$$

so $S_1(u) = o_p(1)$. Lemma 5 shows that this convergence holds uniformly over $u \in \mathcal{U}$.

Consider $S_2(u)$. Let

$$K_g = K\sqrt{\log(N_g)/N_g} \tag{25}$$

for some sufficiently large constant $K > 0$. Let D_G be the event that

$$\max_{g=1,\dots,G}\sup_{u\in\mathcal{U}} \|\hat{\delta}_g(u) - \delta_g(u)\| \leq K_g,$$

and let D_G^c be the event that D_G does not hold. Using Theorem 1, choose a constant K sufficiently large so that

$$P\left(\sup_{u\in\mathcal{U}}\|\hat{\delta}_g(u) - \delta_g(u)\| > K_g\right) \leq CN_g^{-3}$$

for some $C > 0$. By the union bound, $P(D_G^c) \leq CGN_g^{-3}$. By Assumption 9, $CGN_g{}^{3} \to 0$. Therefore,

$$S_2(u) = S_2(u)1\{D_G\} + S_2(u)1\{D_G^c\} = S_2(u)1\{D_G\} + o_p(1)$$

uniformly over $u \in \mathcal{U}$. Further, $\|S_2(u)\|1\{D_G\} \leq C\sum_{g=1}^{G}(r_{1,g} + r_{2,g} + r_{3,g})/\sqrt{GN_g}$ where

$$\begin{aligned} r_{1,g} &= \sup_{u\in\mathcal{U}}\sup_{\delta\in\mathbb{R}^{d_z+1}:\|\delta-\delta_g(u)\|\leq K_g}\left\|J_g^{-1}(u)(\mathbb{G}^g(h_{\delta,u}) - \mathbb{G}^g(h_{\delta_g(u),u}))\right\|\,\|w_g\|, \\ r_{2,g} &= \sup_{u\in\mathcal{U}}\left\|J_g^{-1}(u)\sum_{i=1}^{N_g} h_{\hat{\delta}_g(u),u}(z_{ig}, y_{ig})/\sqrt{N_g}\right\|\,\|w_g\|, \\ r_{3,g} &= \sup_{u\in\mathcal{U}}\sup_{\delta\in\mathbb{R}^{d_z+1}:\|\delta-\delta_g(u)\|\leq K_g}\left\|E_g\left[\sqrt{N_g}(J_g^{-1}(u)h_{\delta,u}(z_{ig}, y_{ig}) - (\delta - \delta_g(u)))\right]\right\|\,\|w_g\|. \end{aligned}$$

By Lemma 4, $E[r_{1,g}] \leq C \log(N_g)/\sqrt{N_g}$ for some $C > 0$. By the optimality of $\hat{\delta}_g(u)$ and since y_{ig} has a continuous conditional distribution, $r_{2,g} \leq C\|w_g\|/\sqrt{N_g}$, and so $E[r_{2,g}] \leq C/\sqrt{N_g}$ for some $C > 0$. Finally, by Lemma 1,

$$E[r_{3,g}] \leq C\sqrt{N_g}K_g^2 \leq C\log(N_g)/\sqrt{N_g}$$

for some $C > 0$. Hence, by Assumption 8,

$$E\left[\sup_{u\in\mathcal{U}} \|S_2(u)\| 1\{D_G\}\right] \leq C\sqrt{G} \max_{g=1,\ldots,G} (\log(N_g)/N_g) = o(1),$$

implying that $\sqrt{G}(\hat{\beta}(u) - \tilde{\beta}(u)) = o_p(1)$ uniformly over $u \in \mathcal{U}$ and completing the first step.

To prove that $\sqrt{G}(\tilde{\beta}(\cdot) - \beta(\cdot)) \Rightarrow \mathbb{G}(\cdot)$ in $\ell^\infty(\mathcal{U})$ (the second step), note that by Assumption 5,

$$\|w_g(\varepsilon(u_1, \eta_g) - \varepsilon(u_2, \eta_g))\| \leq C\|w_g\||u_1 - u_2|,$$

and by Assumption 3, $E[\|w_g\|^2] \leq C$ for some $C > 0$. Therefore, combining Theorems 2.7.11 and 2.5.6 of Van der Vaart and Wellner (1996) implies that $\sum_{g=1}^G w_g\varepsilon(\cdot, \eta_g)/\sqrt{G} \Rightarrow \tilde{\mathbb{G}}(\cdot)$ in $\ell^\infty(\mathcal{U})$ where $\tilde{\mathbb{G}}(\cdot)$ is the Gaussian process with uniformly continuous sample paths and covariance function

$$E[\tilde{\mathbb{G}}(u_1)\tilde{\mathbb{G}}(u_2)'] = E[\varepsilon(u_1, \eta_g)\varepsilon(u_2, \eta_g)w_g w_g'].$$

Combining this convergence result with that in (24) gives the second step. □

Proof of Theorem 3. Equation (24) in the proof of Theorem 2 gives $\hat{S} \to_p S$. Therefore, it suffices to prove that $\|\hat{J}(u_1, u_2) - J(u_1, u_2)\| = o_p(1)$ uniformly over $u_1, u_2 \in \mathcal{U}$.

Note that $\alpha_g(u) - x_g\beta = \varepsilon(u, \eta_g)$. Hence,

$$\hat{\alpha}_g(u) - x_g'\hat{\beta} = (\hat{\alpha}_g(u) - \alpha_g(u)) - x_g'(\hat{\beta}(u) - \beta(u)) + \varepsilon(u, \eta_g) = I_{1,g}(u) - I_2(u) + \varepsilon(u, \eta_g)$$

where $I_{1,g}(u) = \hat{\alpha}_g(u) - \alpha_g(u)$ and $I_2(u) = x_g'(\hat{\beta}(u) - \beta(u))$. Further, we have

$$\sum_{g=1}^{G}(\varepsilon(u_1, \eta_g)\varepsilon(u_2, \eta_g)w_g w_g')/G \to_p J(u_1, u_2)$$

uniformly over $u_1, u_2 \in \mathcal{U}$ by the uniform law of large numbers with bracketing (Theorem 2.4.1 in

Van der Vaart and Wellner (1996)). In addition,

$$\left\|\sum_{g=1}^{G}(I_{1,g}(u_1)I_{1,g}(u_2)w_g w_g')/G\right\| \leq \left\|\sum_{g=1}^{G}K_g^2 w_g w_g'\right\|/G + o_p(1)$$
$$\leq \max_{g=1,\ldots,G} K_g^2 O_p(1) + o_p(1) = o_p(1)$$

uniformly over $u_1, u_2 \in \mathcal{U}$ as in the proof of Theorem 2 (see equation (25) for the definition of K_g). Further,

$$\left\|\sum_{g=1}^{G}(I_{1,g}(u_1)\varepsilon(u_2,\eta_g)w_g w_g')/G\right\| \leq \sum_{g=1}^{G}K_g\left\|\varepsilon(u_2,\eta_g)w_g w_g'\right\|/G + o_p(1)$$
$$\leq \sum_{g=1}^{G}K_g\left\|\sup_{u\in\mathcal{U}}|\varepsilon(u,\eta_g)|w_g w_g'\right\| + o_p(1) = o_p(1)$$

uniformly over $u_1, u_2 \in \mathcal{U}$ where the last equality follows because it follows from Assumption 5 that

$$E\left[\left\|\sup_{u\in\mathcal{U}}|\varepsilon(u,\eta_g)|w_g w_g'\right\|\right] \leq C$$

uniformly over $u_2 \in \mathcal{U}$ for some $C > 0$. Similarly,

$$\left\|\sum_{g=1}^{G}(I_2(u_1)I_2(u_2)w_g w_g')/G\right\| \leq C\sup_{u\in\mathcal{U}}\|\hat{\beta}(u)-\beta(u)\|^2 O_p(1) = o_p(1),$$
$$\left\|\sum_{g=1}^{G}(I_2(u_1)\varepsilon(u_2,\eta_g)w_g w_g')/G\right\| \leq C\sum_{g=1}^{G}\|\varepsilon(u_2,\eta_g)w_g w_g'\|\sup_{u\in\mathcal{U}}\|\hat{\beta}(u)-\beta(u)\|/G = o_p(1)$$

uniformly over $u_1, u_2 \in \mathcal{U}$. Finally,

$$\left\|\sum_{g=1}^{G}(I_{1,g}(u_1)I_2(u_2)w_g w_g')/G\right\| \leq C\sum_{g=1}^{G}K_g\|w_g w_g'\|\sup_{u\in\mathcal{U}}\|\hat{\beta}(u)-\beta(u)\|/G + o_p(1) = o_p(1)$$

uniformly over $u_1, u_2 \in \mathcal{U}$. Combining these inequalities gives the asserted claim. □

Proof of Theorem . It follows from assumptions on singular values of Q_{xw} and eigenvalues of Q_{ww} and $E[\varepsilon(u,\eta_g)^2 w_g w_g']$ that all eigenvalues of $\mathcal{C}(u,u)$ are bounded away from zero uniformly over $u \in \mathcal{U}$. Combining Theorems 2 and 3 and using the Continuous Mapping Theorem then implies that the statistic T converges weakly to some continuous distribution. Further, the statistic T

calculated over a subsample of size b converges to the same distribution since our theory applies to this statistic (in particular, Assumption 8 clearly holds for the subsample). The rest of the proof consists of applying a standard argument for subsampling consistency; see, for example, Theorem 15.7.1 in Lehmann and Romano (2005). □

3.B Tools Appendix

In Appendix A, we used several results from the Theory of Empirical Processes that are used rarely in econometrics literature. For ease of reference, we describe these results in this section.

A semi-metric $\rho : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_+$ is called Gaussian if it can be defined as

$$\rho(f, g) = \left(E[(G(f) - G(g))^2]\right)^{1/2}$$

where G is a tight, zero-mean, Gaussian random element in $l^\infty(\mathcal{F})$. A semi-metric ρ is called Gaussian-dominated if it is bounded from above by Gaussian metric. In particular, it is known that any semi-metric ρ satisfying

$$\int_0^\infty \sqrt{\log N(\varepsilon, \mathcal{F}, \rho)} d\varepsilon < \infty$$

is Gaussian-dominated; see discussion on page 212 in Van der Vaart and Wellner (1996).

Theorem 5 (Bracketing by Gaussian Hypotheses). *For each n, let $Z_{n1}, ..., Z_{nm_n}$ be independent stochastic processes indexed by an arbitrary index set $\mathcal{F}$. Suppose that there exists a Gaussian-dominated semi-metric ρ on $\mathcal{F}$ such that*

$$\begin{aligned}
&(i) \qquad \sum_{i=1}^{m_n} E\left[\|Z_{ni}\|_{\mathcal{F}}\{\|Z_{ni}\|_{\mathcal{F}} > \eta\}\right] \to 0, \text{ for every } \eta > 0,\\
&(ii) \qquad \sum_{i=1}^{m_n} E\left[(Z_{ni}(f) - Z_{ni}(g))^2\right] \leq \rho^2(f, g), \text{ for every } f, g,\\
&(iii) \qquad \sup_{t>0} \sum_{i=1}^{m_n} t^2 P\left(\sup_{f,g \in B(\varepsilon)} |Z_{ni}(f) - Z_{ni}(g)| > t\right) \leq \varepsilon^2,
\end{aligned}$$

for every ρ-ball $B(\varepsilon) \subset \mathcal{F}$ of radius less than ε and for every n. Then the sequence $\sum_{i=1}^{m_n}(Z_{ni} - E[Z_{ni}])$ is asymptotically tight in $l^\infty(\mathcal{F})$. It converges in distribution provided it converges marginally.

Proof. See Theorem 2.11.11 in Van der Vaart and Wellner (1996). ☐

Theorem 6. *Let $\mathcal{F}$ be a P-measurable class of measurable functions with measurable envelope function F. Then*

$$\begin{aligned} (i) \qquad & \|\sup_{f\in\mathcal{F}} |\mathbb{G}_n(f)|\|_{P,p} \le C_p \|J(\theta_n, \mathcal{F})\|F\|_n\|_{P,p}, \\ (ii) \qquad & \|\sup_{f\in\mathcal{F}} |\mathbb{G}_n(f)|\|_{P,p} \le C_p J(1, \mathcal{F})\|F\|_{P,2\vee p},\ 1 \le p. \end{aligned}$$

Here $\theta_n = \sup_{f\in\mathcal{F}} \|f\|_n / \|F\|_n$, where $\|\cdot\|_n$ is the $L_2(P_n)$-semi-norm, and the constant C_p depends only on p.

Proof. See Theorem 2.14.1 in Van der Vaart and Wellner (1996). ☐

Theorem 7. *Let $\mathcal{F}$ be a class of functions $f : \mathcal{X} \to [0,1]$ that satisfies*

$$\sup_Q N(\varepsilon, \mathcal{C}, L_2(Q)) \le \left(\frac{K}{\varepsilon}\right)^V, \text{ for every } 0 < \varepsilon < K$$

where supremum is taken over all probability measures Q. Then for every $t > 0$,

$$P\left(\sup_{f\in\mathcal{F}} |\mathbb{G}_n(f)|\right) \le \left(\frac{Dt}{\sqrt{V}}\right)^V e^{-2t^2}$$

for a constant D that depends on K only.

Proof. See Theorem 2.14.13 in Van der Vaart and Wellner (1996). ☐

Bibliography

Abadie, A., J. Angrist, and G. Imbens, "Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings," *Econometrica*, 2002, *70* (1), 91–117.

Abrevaya, J. and C.M. Dahl, "The effects of birth inputs on birthweight," *Journal of Business and Economic Statistics*, 2008, *26* (4), 379–397.

Allison, Paul D, "Discrete-time methods for the analysis of event histories," *Sociological methodology*, 1982, *13* (1), 61–98.

Alston, R.M., J. R. Kearl, and M.B. Vaughan, "Is There a Consensus among Economists in the 1990s?," *American Economic Review*, 1992, *82*, 203–209.

An Act Enabling Certain Cities and Towns to Control Rents and Evictions, 1970 Mass. Acts 842.

Angrist, J.D. and J. Guryan, "Does teacher testing raise teacher quality? Evidence from state certification requirements," *Economics of Education Review*, 2008, *27* (5), 483–503.

_ **and J.S. Pischke**, *Mostly harmless econometrics: An empiricist's companion*, Princeton Univ Pr, 2009.

_ **and K. Lang**, "Does school integration generate peer effects? Evidence from Boston's Metco Program," *The American Economic Review*, 2004, *94* (5), 1613–1634.

_ **, V. Chernozhukov, and I. Fernández-Val**, "Quantile Regression under Misspecification, with an Application to the US Wage Structure," *Econometrica*, 2006, *74* (2), 539–563.

Arellano, Manuel and Stéphane Bonhomme, "Random Effects Quantile Regression," 2013. Working Paper.

Arnott, Richard, "Time for Revisionism on Rent Control," *Journal of Economic Perspectives*, 1995, *9*, 99–120.

Athey, S. and G. Imbens, "Identification and Inference in Nonlinear Difference-in-Differences Models," *Econometrica*, 2006, *74* (2).

_ **and P.A. Haile**, "Nonparametric approaches to auctions," *Handbook of Econometrics*, 2007, *6*, 3847–3965.

Atlantic Marketing Research, "Impacts of the Termination of Rent Control on Population, Housing Costs and Housing Stock," Technical Report, Atlantic Marketing Research January 1998.

Autor, David H, D. Dorn, and G.H. Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 2013. Forthcoming.

_ **, Lawrence F Katz, and Melissa S Kearney**, "Trends in US wage inequality: Revising the revisionists," *The Review of Economics and Statistics*, 2008, *90* (2), 300–323.

Bajari, Patrick, Chenghuan Sean Chu, and Minjung Park, "An Empirical Model of Subprime Mortgage Default From 2000 to 2007," December 2008. National Bureau of Economic Research Working Paper #14625.

Ballou, D. and M. Podgursky, "The case against teacher certification," *Public Interest*, 1998, pp. 17–29.

Baum-Snow, N., "Did highways cause suburbanization?," *The Quarterly Journal of Economics*, 2007, *122* (2), 775–805.

Belloni, A., V. Chernozhukov, and C. Hansen, "Conditional Quantile Processes Based on Series or Many Regressors," 2006. Working Paper.

Berger, Allen N and Gregory F Udell, "The institutional memory hypothesis and the procyclicality of bank lending behavior," *Journal of Financial Intermediation*, 2004, *13* (4), 458–495.

Bhardwaj, Geetesh and Rajdeep Sengupta, "Subprime Loan Quality," 2011. Federal Reserve Bank of St. Louis Working Paper # 2008-036E.

_ **and** _ , "Subprime mortgage design," *Journal of Banking & Finance*, 2012, *36* (5), 1503–1519.

Bhutta, Neil, Jane Dokko, and Hui Shan, "The Depth of Negative Equity and Mortgage Default Decisions," 2010. Finance and Economics Discussion Series Working Paper #2010-35.

Board of Governors of the Federal Reserve System, "Report to the Congress on Credit Scoring and Its Effects on the Availability and Affordability of Credit," Technical Report, http://www.federalreserve.gov/boarddocs/RptCongress/creditscore/creditscore.pdf August 2007.

Brueckner, Jan K., Paul S. Calem, and Leonard I. Nakamura, "Subprime mortgages and the housing bubble," *Journal of Urban Economics*, 2012, *71* (2), 230 – 243.

Bulow, Jeremy and Paul Klemperer, "Regulated Prices, Rent-Seeking, and Consumer Surplus," *Journal of Economic Perspectives*, 2012, *120(1)*, 160–186.

Caballero, Ricardo and Eduardo Engel, "Explaining Investment Dynamics in U.S. Manufacturing: A Generalized (S,s) Approach," *Econometrica*, 1999, *67(4)*, 783–826.

Campbell, John, Stefano Giglio, and Parag Pathak, "Forced Sales and House Prices," *American Economic Review*, 2011, *101(5)*, 2108–2131.

Canay, I., "A Simple Approach to Quantile Regression for Panel Data," *Econometrics Journal*, 2011, *14*, 368–386.

Cantor, Patricia, "Massachusetts Defeats Rent Control," 1995. National Housing Institute, April/May, Available at: http://www.nhi.org/online/issues/80/massrent.html, Last accessed: December 12, 2011.

Card, David and Alan Krueger, *Myth and Measurement: The New Economics of the Minimum Wage*, Princeton University Press, 1995.

Chamberlain, G., "Quantile regression, censoring, and the structure of wages," in "Advances in Econometrics, Sixth World Congress," Vol. 1 New York: Cambridge University Press 1994, pp. 171–209.

Chernozhukov, Victor and Christian Hansen, "An IV model of quantile treatment effects," *Econometrica*, 2005, *73* (1), 245–261.

_ and _ , "Instrumental quantile regression inference for structural and treatment effect models," *Journal of Econometrics*, 2006, *132* (2), 491–525.

_ and _, "Instrumental variable quantile regression: A robust inference approach," *Journal of Econometrics*, 2008, *142* (1), 379–398.

_, Denis Chetverikov, and Kengo Kato, "Gaussian approximation of suprema of empirical processes," 2012. Working Paper.

_, _, and _, "Anti-concentration and honest, adaptive confidence bands," 2013. Working Paper.

_, _, and _, "Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors," *The Annals of Statistics*, 2013, *41* (6), 2786–2819.

_, I. Fernandez-Val, J. Hahn, and W. Newey, "Average and Quantile Effects in Nonseparable Panel Models," *Econometrica, forthcoming*, 2012, pp. 1–83.

Chesher, Andrew, "Identification in nonseparable models," *Econometrica*, 2003, *71* (5), 1405–1441.

(COP) Congressional Oversight Panel of the Troubled Assets Recovery Program, "Foreclosure Crisis: Working toward a Solution," Technical Report, http://cop.senate.gov/reports/library/report-030609-cop.cfm COP Report, March 6, 2009.

Corbae, Dean and Erwan Quintin, "Leverage and the Foreclosure Crisis," August 2013. National Bureau of Economic Research Working Paper #19323.

Cunningham, Chris and Robert R Reed, "Negative equity and wages," *Regional Science and Urban Economics*, 2013, *43* (6), 817–825.

Dagher, Jihad and Ning Fu, "What Fuels the Boom Drives the Bust: Regulation and the Mortgage Crisis," 2011. IMF Working Paper WP/11/215.

Davis, Lucas and Lutz Killian, "The Allocative Cost of Price Ceilings in the U.S. Residential Market for Natural Gas," *Journal of Political Economy*, 2011, *119(2)*, 212–241.

Dell'Ariccia, Giovanni, Deniz Igan, and Luc Laeven, "Credit booms and lending standards: Evidence from the subprime mortgage market," *Journal of Money, Credit and Banking*, 2012, *44* (2-3), 367–384.

Demyanyk, Yuliya and Otto Van Hemert, "Understanding the subprime mortgage crisis," *Review of Financial Studies*, 2011, *24* (6), 1848–1880.

Deng, Yongheng, John M. Quigley, and Robert van Order, "Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options," *Econometrica*, 2000, *68* (2), pp. 275–307.

Elbers, Chris and Geert Ridder, "True and spurious duration dependence: The identifiability of the proportional hazard model," *The Review of Economic Studies*, 1982, *49* (3), 403–409.

Ellison, Glenn, Edward Glaeser, and William Kerr, "What Causes Industry Agglomeration? Evidence from Co-agglomeration Patterns," *American Economic Review*, 2010, *100(3)*, 1191–1213.

Epple, Dennis, "Rent Control with Reputation: Theory and Evidence," *Regional Science and Urban Economics*, 1988, *28(6)*, 679–710.

Fain, Howard, "Spotlight: Cambridge and Proportional Representation," *Voting and Democracy*, 1995.

Fallis, George and Lawrence B. Smith, "Uncontrolled Prices in a Controlled Market: The Case of Rent Controls," *American Economic Review*, 1984, *74*, 193–200.

Feenstra, Robert C, *Offshoring in the global economy: microeconomic structure and macroeconomic implications*, Cambridge Univ Press, 2010.

– **and Gordon H Hanson**, "The impact of outsourcing and high-technology capital on wages: estimates for the United States, 1979–1990," *The Quarterly Journal of Economics*, 1999, *114* (3), 907–940.

Feldstein, Martin S., "How to Stop the Drop in Home Values," *The New York Times*, October 12, 2011.

Fetter, Daniel, "The Home Front: Rent Control and the Rapid Wartime Increase in Home Ownership," 2013. Unpublished paper, Wellesley College.

Finkel, Meryl and James E. Wallace, "Cambridge Housing Survey," 1987. Abt Associates.

Foote, Christopher L, Kristopher S Gerardi, and Paul S Willen, "Negative equity and foreclosure: Theory and evidence," *Journal of Urban Economics*, 2008, *64* (2), 234–245.

– , – , **Lorenz Goette, and Paul S Willen**, "Reducing Foreclosures: No Easy Answers," in Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, eds., *NBER Macroeconomics Annual 2009, Volume 24*, University of Chicago Press, April 2010, pp. 89–138.

Frankena, Mark, "Alternative models of rent control," *Urban Studies*, 1975, *12*, 303–308.

Frech, H.E. and William Lee, "The Welfare Cost of Rationing-by-Queuing across Markets: Theory and Estimates from the U.S. Gasoline Crises," *Quarterly Journal of Economics*, 1987, *102(1)*, 97–108.

Friedman, Milton and George Stigler, "Roofs or Ceilings? The Current Housing Problem," in "Popular Essays on Current Problems," Vol. 1(2), Irvington-on-Hudson, New York: Foundation for Economic Education, 1946.

Fujita, Masahisa, *Urban Economic Theory: Land Use and City Size*, Cambridge University Press, 1991.

Galvao, A.F., "Quantile regression for dynamic panel data with fixed effects," *Journal of Econometrics*, 2011.

_ **and L. Wang**, "Efficient minimum distance estimator for quantile regression fixed effects panel data.," 2013. Working Paper.

Genesove, David and Chris Mayer, "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *Quarterly Journal of Economics*, 2001, pp. 1233–1260.

_ **and Christopher J Mayer**, "Equity and Time to Sale in the Real Estate Market," *The American Economic Review*, 1997, pp. 255–269.

Gerardi, Kris, Adam Shapiro, and Paul Willen, "Subprime Outcomes: Risky Mortgages, Homeownership Experiences, and Foreclosures," 2007. FRB Boston Working Paper Series, paper no. 07-15.

Gerardi, Kristopher, Adam Shapiro, and Paul Willen, "Subprime outcomes: Risky mortgages, homeownership experiences, and foreclosures," 2007. FRB of Boston Working Paper No. 07-15.

_ **, Andreas Lehnert, Shane M Sherlund, and Paul Willen**, "Making sense of the subprime crisis," *Brookings Papers on Economic Activity*, 2008, (2), 69–159.

Glaeser, Edward, *Cities, Agglomeration and Spatial Equilibrium*, Oxford University Press, 2008.

_ **and Erzo Luttmer**, "The Misallocation of Housing under Rent-Control," *American Economic Review*, 2003, *93(4)*, 1027–1046.

– **and Joseph Gyourko**, *Rethinking Federal Housing Policy: How to Make Housing Plentiful and Affordable*, AEI Press, 2009.

Glaeser, Edward L., *Cities, Agglomeration, and Spatial Equilibrium* The Lindahl Lectures, Oxford: Oxford University Press, 2008.

Government Accountability Office, "Nonprime Mortgages: Analysis of Loan Performance, Factors Associated with Defaults, and Data Sources," Technical Report 2010.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti, "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," *Journal of Political Economy*, 2010, *118(3)*, 536–598.

Gyourko, Joseph and Peter Linneman, "Equity and Efficiency Aspects of Rent Control: An Empirical Study of New York City," *Journal of Urban Economics*, 1989, *26(1)*, 54–74.

– **and** – , "Rent Controls and Rental Housing Quality: A Note on the Effects of New York City's Old Controls," *Journal of Urban Eonomics*, 1990, *27*, 398–409.

Hahn, Jinyong and Juergen Meinecke, "Time-invariant regressor in nonlinear panel model with fixed effects," *Econometric Theory*, 2005, *21* (02), 455–469.

Haile, P., H. Hong, and M. Shum, "Nonparametric tests for common values at first-price sealed-bid auctions," 2003. Working Paper, Yale.

Han, Aaron and Jerry A Hausman, "Flexible parametric estimation of duration and competing risk models," *Journal of Applied Econometrics*, 1990, *5* (1), 1–28.

Hartley, Dan, "The Effect of Foreclosures on Nearby Housing Prices: Supply or Disamenity?," 2010. FRB of Cleveland Working Paper No. 10-11.

Haskel, Jonathan, Robert Z Lawrence, Edward E Leamer, and Matthew J Slaughter, "Globalization and US wages: Modifying classic theory to explain recent facts," *The Journal of Economic Perspectives*, 2012, pp. 119–139.

Haughwout, Andrew, Richard Peach, and Joseph Tracy, "Juvenile delinquent mortgages: Bad credit or bad economy?," *Journal of Urban Economics*, 2008, *64* (2), 246 – 257.

Hausman, Jerry A, "Mismeasured variables in econometric analysis: Problems from the right and problems from the left," *The Journal of Economic Perspectives*, 2001, *15* (4), 57–67.

_ **and Tiemen M Woutersen**, "Estimating a semi-parametric duration model without specifying heterogeneity," 2010. Working Paper, The Johns Hopkins University, Department of Economics.

_ **and W. E. Taylor**, "Panel data and unobservable individual effects," *Econometrica*, 1981, pp. 1377–1398.

_ **, Ye Luo, and Christopher Palmer**, "Errors in the dependent variable of conditional quantile models," 2014. Working Paper.

Haveman, Judith, "Mass. City Gets a New Lease on Life; Cambridge Booms as Rent Control Ends, but the Poorer Pay a Price," 1998. Washington Post, September 19, A1.

Heckman, James J and Bo E Honoré, "The identifiability of the competing risks model," *Biometrika*, 1989, *76* (2), 325–330.

_ **and Burton Singer**, "A method for minimizing the impact of distributional assumptions in econometric models for duration data," *Econometrica: Journal of the Econometric Society*, 1984, pp. 271–320.

_ **and Richard Robb**, "Alternative methods for evaluating the impact of interventions: An overview," *Journal of Econometrics*, 1985, *30* (1), 239–267.

Holzer, Harry, Lawrence Katz, and Alan Krueger, "Job Queues and Wages," *Quarterly Journal of Economics*, 1991, *106(3)*, 739–768.

Horowitz, Joel L, "Semiparametric estimation of a proportional hazard model with unobserved heterogeneity," *Econometrica*, 1999, *67* (5), 1001–1028.

Hubbard, R Glenn and Christopher J Mayer, "The Mortgage Market Meltdown and House Prices," *The B.E. Journal of Economic Analysis & Policy*, 2009, *9* (3).

Hurst, Erik, Veronica Guerrieri, and Dan Hartley, "Endogenous Gentrification and Housing Price Dynamics," 2013. *Journal of Public Economics*, forthcoming.

Imbens, Guido and Jeffrey M Wooldridge, "Control Function and Related Methods," What's New in Econometrics, Summer 2007.

Imbens, Guido W, "Nonadditive Models with Endogenous Regressors," in R. Blundell, W. Newey, and T. Persson, eds., *Advances in Economics and Econometrics*, Vol. III, Cambridge University Press, 2007, chapter 2, pp. 17–46.

_ **and Whitney K Newey**, "Identification and estimation of triangular simultaneous equations models without additivity," *Econometrica*, 2009, *77* (5), 1481–1512.

Ioannides, Yannis, "Interactive Property Valuations," *Journal of Urban Economics*, 2003, *53*, 145–170.

(JEC) Joint Economic Committee, "The Subprime Lending Crisis: The Economic Impact on Wealth, Property Values and Tax Revenues, and How We Got Here," Technical Report October 2007.

Kasy, Max, "Identification in a Model of Sorting with Social Externalities and the Causes of Urban Segregation," 2013. Unpublished working paper, Harvard University.

Kato, K., A.F. Galvao, and G. Montes-Rojas, "Asymptotics for Panel Quantile Regression Models with Individual Effects," *Journal of Econometrics*, 2011, *forthcoming.*

_ **and** _ , "Smoothed Quantile Regression for Panel Data," *working paper*, 2011.

Katz, Lawrence F and David H Autor, "Changes in the wage structure and earnings inequality," *Handbook of labor economics*, 1999, *3*, 1463–1555.

Kau, James B., Donald C. Keenan, Constantine Lyubimov, and V. Carlos Slawson, "Subprime mortgage default," *Journal of Urban Economics*, 2011, *70* (2–3), 75 – 87.

Keys, Benjamin J, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig, "Did securitization lead to lax screening? Evidence from subprime loans," *The Quarterly Journal of Economics*, 2010, *125* (1), 307–362.

Koenker, R., "Quantile regression for longitudinal data," *Journal of Multivariate Analysis*, 2004, *91* (1), 74–89.

_ **and G. Bassett Jr**, "Regression quantiles," *Econometrica*, 1978, pp. 33–50.

Krainer, John and Elizabeth Laderman, "Prepayment and delinquency in the mortgage crisis period," September 2011. Federal Reserve Bank of San Francisco Working Paper 2011-25.

Krugman, Paul, "Technology, trade and factor prices," *Journal of International Economics*, 2000, *50* (1), 51–71.

_ , "A Catastrophe Foretold," *The New York Times*, October 26, 2007.

_ , "Some Housing Pictures," *The New York Times*, October 27, 2007.

Lamarche, C., "Robust penalized quantile regression estimation for panel data," *Journal of Econometrics*, 2010, *157* (2), 396–408.

Lancaster, Tony, "Generalised residuals and heterogeneous duration models: With applications to the Weilbull model," *Journal of Econometrics*, 1985, *28* (1), 155–169.

Landvoigt, Tim, Monika Piazzesi, and Martin Schneider, "The Housing Market(s) of San Diego," 2012. Stanford University, working paper.

Larsen, B., "The Efficieny of Dynamic, Post-Auction Bargaining: Evidence from Wholesale Used-Auto Auctions," 2012. Working Paper, MIT.

_ , "Occupational Licensing and Quality: Distributional and Heterogeneous Effects in the Teaching Profession," 2012. Working Paper, MIT.

Leamer, Edward E, "Trade, wages and revolving door ideas," 1994. National Bureau of Economic Research Working Paper # 4716.

Lee, David and Emmauel Saez, "Optimal Minimum Wages in Competitive Labor Markets," 2010. NBER Working Paper, No. 14320.

Lee, Sokbae, "Endogeneity in quantile regression models: A control function approach," *Journal of Econometrics*, 2007, *141* (2), 1131–1158.

Lehmann, E. and J. Romano, *Testing Statistical Hypotheses*, Springer Texts in Statistics, 2005.

Leonard, Dutch, "Housing Regulation in Cambridge: Its Goals and Effects," 1981. Cambridge Chamber of Commerce.

Linden, Leigh and Jonah Rockoff, "There Goes the Neighborhood? Estimates of the Impacts of Crime Risk on Property Values from Megan's Law," *American Economic Review*, 2008, *98(3)*, 1103–1127.

Mayer, Christopher J., "Comment on "Reducing Foreclosures: New Easy Answers"," in Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, eds., *NBER Macroeconomics Annual 2009, Volume 24*, University of Chicago Press, April 2010, pp. 139–148.

_ **and Karen Pence**, "Subprime mortgages: what, where, and to whom?," 2008. National Bureau of Economic Research Working Paper #National Bureau of Economic Research Working Paper #14083.

_ **and Todd M. Sinai**, "U.S. House Price Dynamics and Behavioral Finance," in Christopher L Foote, Lorenz Goette, and Stephan Meier, eds., *Policymaking Insights from Behavioral Economics*, Boston: Federal Reserve Bank of Boston, 2009, pp. 259–297.

_ **, Karen Pence, and Shane M Sherlund**, "The rise in mortgage defaults," *Journal of Economic Perspectives*, 2009, *23* (1), 27–50.

_ **, Tomasz Piskorski, and Alexei Tchistyi**, "The Inefficiency of Refinancing: Why Prepayment Penalties Are Good for Risky Borrowers," December 2010. National Bureau of Economic Research Working Paper #16586.

McCall, Brian P, "Unemployment insurance rules, joblessness, and part-time work," *Econometrica*, 1996, pp. 647–682.

Meyer, Bruce D, "Unemployment Insurance and Unemployment Spells," *Econometrica*, 1990, *58* (4), 757–782.

Mian, Atif, "Comment on "Reducing Foreclosures: No Easy Answers"," in Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, eds., *NBER Macroeconomics Annual 2009, Volume 24*, University of Chicago Press, April 2010, pp. 149–156.

_ **, Amir Sufi, and Francesco Trebbi**, "Foreclosures, House Prices, and the Real Economy," January 2011. National Bureau of Economic Research Working Paper #16685.

_ **and** _ , "The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis," *Quarterly Journal of Economics*, 2009, *124(4)*, 1449–1496.

_ **and** _ , "What explains high unemployment? The aggregate demand channel," 2012. National Bureau of Economic Research Working Paper #17830.

Milgrom, Paul and John Roberts, "Rationalizability, Learning and Equilibrium in Games With Strategic Complementarities," *Econometrica*, 1990, *58*, 1255–1278.

Murphy, Kevin M. and Robert H. Topel, "Estimation and Inference in Two-Step Econometric Models," *Journal of Business & Economic Statistics*, 1985, *3* (4), 370–79.

Olsen, Edgar, "An Econometric Analysis of Rent Control," *Journal of Political Economy*, 1972, *86(1)*, 1081–1100.

Oser, Alan S., "Ending Rent Controls the Massachusetts Way," 1996. *New York Times*, April 21st.

Pagan, Adrian, "Econometric issues in the analysis of regressions with generated regressors," *International Economic Review*, 1984, *25* (1), 221–247.

Palmer, C., "Suburbanization and Urban Decline," 2011. Working Paper, MIT.

Pavlov, Andrey and Susan Wachter, "Subprime Lending and House Price Volatility," 2009. U of Penn, Inst for Law & Econ Research Paper 08-33.

Pennington-Cross, Anthony and Giang Ho, "The Termination of Subprime Hybrid and Fixed-Rate Mortgages," *Real Estate Economics*, 2010, *38* (3), 399–426.

Pollakowski, Henry, "Rent Control and Housing Investment: Evidence from Deregulation in Cambridge, Massachusetts," 2003. Manhattan Institute Civic Report, May.

Pope, Jaren C., "Fear of Crime and Housing Prices: Household Reactions to Sex Offender Registries," *Journal of Urban Economics*, November 2008, *64* (3), 601–614.

Prentice, Ross L and Lynn A Gloeckler, "Regression analysis of grouped survival data with application to breast cancer data," *Biometrics*, 1978, pp. 57–67.

Rent Control Board, 1982. Archives of the Cambridge Historical Commission.

Rivers, Douglas and Quang H Vuong, "Limited information estimators and exogeneity tests for simultaneous probit models," *Journal of Econometrics*, 1988, *39* (3), 347–366.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Raymond Owens, "Housing Externalities," *Journal of Political Economy*, 2010, *118(3)*, 829–858.

Schwartz, Amy Ellen, Ingrid Gould Ellen, Ioan Voicu, and Michael Schill, "The External Effects of Place-Based Subsidized Housing," *Journal of Urban Economics*, 2006, *36(6)*, 679–707.

Sherlund, Shane M, "The Past, Present, and Future of Subprime Mortgages," 2008. Finance and Economics Discussion Series Working Paper #2008-63.

Sims, David P., "Out of Control: What can we learn from the end of Massachusetts rent control?," *Journal of Urban Economics*, 2007, *61(1)*, 129–151.

_ , "Rent Control Rationing and Community Composition: Evidence from Massachusetts," *The B.E. Journal of Economic Analysis & Policy*, 2011, *11(1)*.

Sinai, Todd M., "House Price Moments in Boom-Bust Cycles," May 2012. National Bureau of Economic Research Working Paper #18059.

Strange, William, "Overlapping Neighborhoods and Housing Externalities," *Journal of Urban Economics*, 1992, *32*, 17–39.

Suen, Wing, "Rationing and Rent Dissipation in the Presence of Heterogeneous Individuals," *Journal of Political Economy*, 1989, *97(6)*, 1384–1394.

Tolbert, Charles M and Molly Sizer, "US commuting zones and labor market areas: A 1990 update," 1996. Economic Research Service Staff Paper No. 9614.

Tucker, William, "Anarchy, State and Rent Control," 1986. *The New Republic*, December 22, pp. 20-21.

Van der Vaart, A.W. and J.A Wellner, *Weak Convergence and Empirical Processes*, Springer Series in Statistics, 1996.

Wang, Shing-Yi, "State Misallocation and Housing Prices: Theory and Evidence from China," *American Economic Review*, 2011, *101(5)*, 2081–2107.

Wheaton, William C., "A Proposal to Repair US Residential Mortgages." Unpublished mimeo, MIT.

White, H., "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity," *Econometrica*, 1980, pp. 817–838.

Wooldridge, Jeffrey M, *Econometric analysis of cross section and panel data*, The MIT press, 2002.