Great Ideas (Don't) Sell Themselves: The Disclosure Paradox in Digital Startups Auctions

By

Luca Gius

B.S. in Economics
University of Pisa and S. Anna, 2016

M.S. in Economics
University of Pisa, 2018

SUBMITTED TO THE DEPARTMENT OF MANAGEMENT IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN MANAGEMENT RESEARCH

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2023

Authored by: Luca Gius

MIT Sloan School of Management

January 20, 2023

Certified by: Scott Stern

Professor of MIT Sloan School of Management, Thesis Supervisor

Accepted by: Eric So
Sloan Distinguished Professor of Financial
Economics Professor, Accounting and Finance
Faculty Chair MIT Sloan PhD Program

Great Ideas (Don't) Sell Themselves:
The Disclosure Paradox in Digital Startups Auctions

by

Luca Gius

Submitted to the Department of Management
on January 20, 2023, in partial fulfillment of the
requirements for the degree of
Master of Science in Management Research

Abstract

An inventor may be hesitant to reveal her idea to potential buyers because she fears it will be stolen. I show that this disclosure paradox can result in only the worst ideas being sold, leading to an ineffective commercialization of the best ideas. This distortion at the top is especially significant when gains from trade are concentrated among a few, high-quality ideas.

Next, I investigate the disclosure paradox in an online marketplace for early-stage digital start-ups. Sellers can protect their unpatentable ideas through confidentiality agreements (NDAs), and they can have the exchange verify their advertising revenues. I provide evidence that these disclosure technologies are used by high-quality inventors to assuage information frictions, sell their ideas more easily and to capture more value. Surprisingly, confidentiality agreements discourage many potential buyers and are less effective than revenue verification.

I estimate that the disclosure paradox leads to significant welfare losses: going from a scenario in which every seller has access to revenue verification to one where nobody has access to it destroys 63% of the potential gains from trade and reduces the number of sold start-ups by 17%. Through simulations, I show that the losses are magnified by the skewed distribution of ideas.

Thesis Supervisor: Scott Stern
Title: David Sarnoff Professor of Management of Technology

I am grateful to (especially) Scott Stern, Sandro Montresor (discussant), Colleen Cunningham (discussant), Pierre Azoulay, Mercedes Delgado, Danielle Li, Eric Von Hippel, participants at the MIT TIES Reading group, at the Druid '21 Conference and the '22 SEFI Workshop for helpful feedback on this project.

Contents

List of Figures

List of Tables

1 Introduction

Inventors typically keep their unpatented ideas under wraps, and rightly so: the simple disclosure of the information underlying an unprotected innovation can result in theft (Arrow [1962]'s "disclosure paradox"). For example, Robert Kearns, the inventor of the intermittent windshield wiper, was sidelined by Ford Motor Company in the commercialization process; Kearns was recognized as the legitimate inventor only after a 20-year long lawsuit.[1]

The potential for expropriation can reduce the economic incentives for technological innovation (Nelson [1959], Arrow [1962]). More subtly, Arrow's disclosure problem has the potential to stifle the exchange of ideas through markets (Anton and Yao [1994], Arora and Gambardella [2010b], Gans and Stern [2003]) and the division of innovative labor (Arora and Gambardella [2010b]). As emphasized by Teece [1986], the organization that is most effective at commercializing an innovation may be different from the one that develops the initial idea. In an inefficient market for technology, this can prevent valuable ideas from being developed into fully-fledged products.

The primary purpose of this paper is to investigate when the "disclosure paradox" leads to sizable welfare losses. This critically depends on what ideas remain unsold when disclosure is limited, and whether those ideas are significantly harder to commercialize unless traded to a more effective owner (i.e.: the size of the lost gains from trade). If, for instance, the disclosure problem resulted in a fixed transaction cost it would disproportionately hamper the sale of innovations with low gains from trade and thus likely be of modest empirical importance. In practice, this is not the case: under realistic assumptions, I show that limited disclosure disproportionately prevents the sale of ideas that are highly valuable and characterized by the highest gains from trade. Then I calibrate the model on bidding data from an online marketplace of digital start-ups, and show that losses due to limited disclosure are sizable and concentrated among a few highly valuable "missed deals": my focus on welfare is a novel contribution to a long-standing literature on the disclosure paradox which has primarily focused either on testing for the existence of adverse selection[2] or on how specific tools affect the timing and success of an idea sale.[3]

In section 2, I begin by developing a simple model of a market for ideas. Inventors have the choice to either commercialize their idea independently (and so minimize the problem of ex ante disclosure) or sell to potential buyers (who are more effective at commercializing the idea). While the quality of the innovation is observable to each inventor, quality is only observable to

[1] https://www.newyorker.com/magazine/1993/01/11/the-flash-of-genius

[2] Most studies have focused on the pharmaceutical industries where patenting, contrarily to the present setting, is widespread: some (Pisano [1997], Guedj [2005]) find that licensed chemical compounds are more likely to fail than internally developed ones. Others (Danzon et al. [2005], Arora et al. [2009]) do not to find evidence of a lemons problem.

[3] E.g. patent publication and grant (Gans et al. [2008], Hegde and Luo [2017]), biomedical licensing contracts (Hegde [2014]), observable proxies for idea quality and seller reputation (e.g.: movie industry experience as in Luo [2014]), and venture capital certification (Megginson and Weiss [1991], Hsu [2006]).

potential buyers if the inventors are willing to disclose the underlying details of their inventions. However, in the absence of full intellectual property rights, potential buyers can expropriate ideas, motivating the sellers to limit disclosure. Lack of disclosure causes classic asymmetric information issues (Akerlof [1970]), Levin [2001]), whereby buyers cannot tell good ideas and bad ideas apart. Framing Arrow's disclosure paradox as an asymmetric information issue clarifies that the best ideas are the hardest to sell, and that the welfare losses are large when gains from trade are concentrated among the high quality ideas. This is true, for example, if gains from trade grow with idea quality (e.g.: if bad ideas fail independently from the manager) and the distribution of idea quality is skewed. Worryingly, a rich empirical literature has shown that this is a realistic scenario (Pakes [1986], Harhoff et al. [1999], Giuri et al. [2007], Azevedo et al. [2020]).

To test the model's predictions[4] (formalized in section 3), in section 4 I focus on an online marketplace in which sellers and buyers deal in early-stage digital businesses, ranging from ecommerce stores to SaaS to apps. Sellers cannot protect their business ideas via patents, but they can have their past advertising revenues verified by the marketplace and may require potential buyers to sign confidentiality agreements. The auction data include bids for both successful and unsuccessful transactions, which I can leverage to estimate demand and supply for the start-ups on sale with and without NDAs and revenue verification. This allows me to simulate welfare losses under various scenarios of imperfect information.

While both "disclosure technologies" reduce the incidence of fraud on the platform (by 12% for verification and 4% for NDAs), only revenue verification is consistently associated with better outcomes. Revenue-verified (but not NDA-requesting) projects are characterized by 3x higher bids and allow buyers to capture more value for their ideas (relatively to the initial asking price). Revenue-verified (but not NDA-requesting) businesses are also 7% more likely to sell to be sold. A test in the spirit of (Pisano [1997], Guedj [2005], Danzon et al. [2005], Arora et al. [2009]) suggests that verification effectively reduces asymmetric information[5]: while unverified ideas that are sold are more likely to fail (because people are more likely to sell flawed ideas, or "lemons"), this is not the case if the start-ups' advertising revenues are properly vetted during

[4]In an extension of the model, I allow some sellers to invest in disclosure technologies (e.g: non-disclosure agreements and third-party certification) which reduce buyers' uncertainty about idea quality. I show that better inventors are more likely to disclose, whereas worse inventors refuse to disclose if they can get away with it; complete "unraveling" and full disclosure of all private information (Grossman [1981], Jovanovic [1982], Milgrom [1981], Lewis [2011]) only happens if the disclosure technologies are free and universally available.

[5]This is corroborated by two other tests. First, in a field experiment on Mechanical Turk 291 subjects are randomly exposed to one of two versions of a real auction listing from the data. Those who are shown the version with third-party verified revenues report higher valuations than those exposed to the same listing without certification. Second, Arora and Gambardella [2010a] suggest that if buyers are uncertain about the (common) value of an idea they are competing for, they will place bids lower than their assessment especially as the number of competitors increases. Failure to do so might result in a "winner's curse" (Milgrom and Weber [1982]), whereby the winning bid is too optimistic and ends up overpaying. I show that auctions over disclosed ideas suffer relatively less from the "winner's curse" (i.e: the elasticity of the final bid relatively to the number of bidders is higher), which implies that buyers are less afraid to overpay for disclosed ideas. This is consistent with the notion that disclosure is informative.

the transaction process.

Despite the potential benefits, I observe that only a minority of sellers verifies their revenues (11%) or asks buyers to sign a non-disclosure agreement (5%). In line with the model, this can be explained by the fact that neither technology is both cost-free and universally available. The rare use of NDA agreements, a type of contract in principle ideal to solve disclosure issues, can be traced back to buyer aversion. I document that sellers who ask for non-disclosure agreements receive 25% fewer bids and are 50% more likely not to receive any bid, a sign that buyers are wary to sign NDAs. This might explain why, at least in this context, asymmetric information concerns cannot be completely solved via contracts (Arora and Gambardella [2010b]). A calibration exercise suggests that only extreme buyer aversion (a drop in 90% or more of the likelihood to sell) can justify such a limited use of NDAs. Revenue verification, on the other hand, is cost-free but not universally available: only few sellers have advertising revenues recorded via Google Adsense (as opposed to no advertising revenues, or revenues recorded via other services), which are the only type that the marketplace can verify. Not only does this prevent most sellers from verifying their revenues, but it also allows low-quality sellers to avoid verification without incurring the reputational losses. To show this, I use data from Builtwith.com on the digital infrastructure of the startups' website before the auction went live: I show that not all websites with ads technology decided to have their advertising revenues verified by the marketplace, and that those that chose not to do so are more likely to be lemons (i.e: fail after the auction).

To conclude, I estimate welfare losses due to the disclosure paradox: shutting down free and universally available revenue verification destroys 63% of the potential gains from trade, despite only reducing the number of traded ideas by 17%. Simulations show that this reflects the fat-tailed nature of ideas: the disclosure paradox disproportionately affects a few, particularly important transactions. Shutting down verification when the idea distribution is one third as unequal as the empirical one would result in a loss of only 10% of the potential gains from trade.

2 Theory

The model, both in its simple and more general version, describes a market exchange of ideas. Sellers are unable to disclose their ideas prior to the transaction: buyers first choose a price and then blindly pick one of the ideas which the inventors are willing to sell at that price. Intuitively, the best inventors might refuse to sell at a price reflecting average (as opposed to best) quality. This, in turn, decreases the expected quality of ideas on sale: price and expected quality drop until an equilibrium is reached (Akerlof [1970]).

The key result can be summarized as follows: the more gains from trade are concentrated in the few, good ideas, the more gains will be lost.

2.1 Adverse selection in markets for ideas: a simple example

$L + P$ ideas are up for sale. P ideas ("peaches") are worth b to the buyer and s to the seller, while the other L ideas are worthless ("lemons"). Gains from trade exist ($b > s > 0$), for example because the buyer is better than the seller at commercializing the good ideas.

By assumption, no inventor can credibly disclose the idea's quality without giving the idea away, so there is asymmetric information between buyer and sellers. The buyer picks one idea at random and makes a take-it-or-leave-it offer.

Lemons will sell at any price $p \geq 0$, while peaches only accept offers $p \geq s$. Knowing this, the buyer's expected profit will be:

$$\Pi_{\text{buyer}}(p) = \begin{cases} -p & \text{if} \quad p < s \\ \\ b\left(\frac{P}{L+P}\right) - p & \text{if} \quad p \geq s \end{cases}$$

In equilibrium, the buyer will offer s if $\Pi_{\text{buyer}}(s) \geq 0$. Otherwise, buyers will offer $p = 0$ and only lemons will be sold. This breakdown happens whenever:

$$\frac{b}{s} < \frac{L+P}{P} \tag{1}$$

In this case the lost gains from trade (relative to a perfect information benchmark) are:

$$\mathcal{L} = \frac{(b-s)\left(\frac{P}{L+P}\right)}{b\left(\frac{P}{L+P}\right)} = 1 - \frac{s}{b} \tag{2}$$

Two conditions are necessary for Arrow's disclosure paradox to be especially harmful ($\mathcal{L} \to 100\%$). First, the market needs to break down (condition 1), which happens if the lemon-to-peaches ratio is high enough relatively to the gains from trading the peaches. Second, gains from trade need to be sizable (eq. 2). In other words, welfare losses are bigger when gains from trade are more concentrated among the best ideas.

2.2 The general model

Inventors and potential buyers meet at the marketplace of ideas. Inventors have ideas of privately known quality $\theta \in \Theta$. Potential buyers only know the distribution of quality $F \in \Delta(\Theta)$. Gains from trade exist: sellers value their ideas $s(\theta)$, and buyers value them $b(\theta) \geq s(\theta)$.[6,7]

[6] Potential buyers are managers, acquisition entrepreneurs or incumbents good at developing and commercializing inventions. Buyers have typically access to a range of complementary assets (Teece [1986]) – manufacturing, logistics, marketing channels – which inventors lack for. Åstebro and Serrano [2015] show that partnering with an incumbent firm doubles a startup's commercialization chances and increases revenues by a third. Alternatively, the same disclosure paradox arises when inventors need to reveal their ideas to potential partners and consultants, whose expertise is needed to better screen or commercialize the idea (Biais and Perotti [2008]).

[7] Both $b(.)$ and $s(.)$ are weakly increasing in θ and linear in price. I assume $\mathbb{E}_\theta[b(\theta)] < \infty$. In some of the results I will explicitly assume that $b' \geq s'$ to capture the notion that gains from trade are concentrated among the

Given a posted[8] price p, inventors will sell if and only if:

$$s(\theta) \leq p \tag{3}$$

Buyers will buy at p if and only if:

$$\mathbb{E}_\theta[b(\theta)|s(\theta) \leq p] \geq p \tag{4}$$

Under perfect competition among buyers, the equilibrium price(s) p^* satisfies $p^* = \mathbb{E}_\theta[b(\theta)|s(\theta) \leq p^*]$: the price reflects the average quality of ideas sold *in equilibrium*. Figure 1 shows an example: all ideas above the 66^{th} percentile remain unsold, whereas all ideas below are sold at $p^* = 4.13$. The lost gains from trade are represented by the red dotted area.

It is straightforward to show that p^* always exists and may not be unique: as well known in the literature (Mas-Colell et al. [1995]) this may happen if the demand schedule (specified in 4) is upward sloping. In general, I will take as price equilibrium the pareto-efficient supremum p^* of the set of all price equilibria[9], which always exists and is trivially bounded above by $\mathbb{E}[b(\theta)]$.

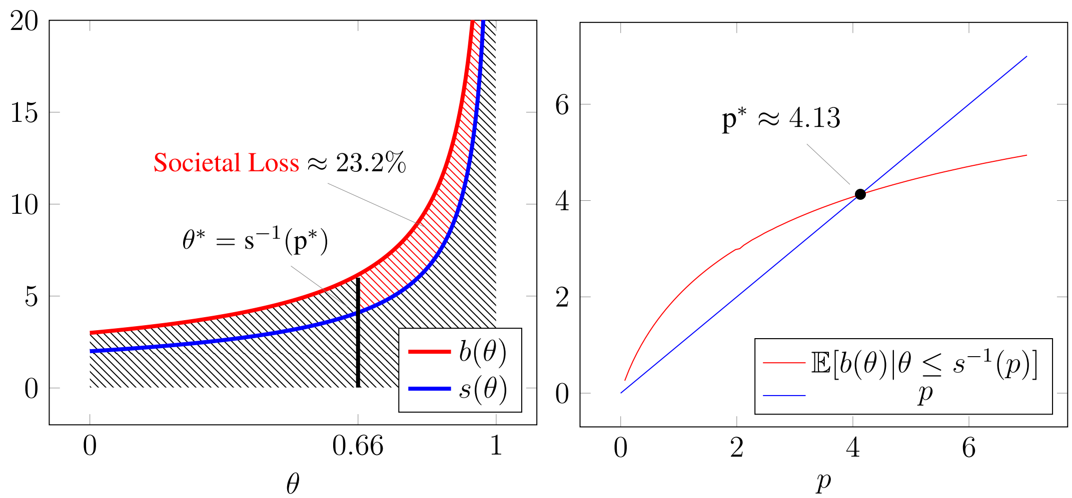


Figure 1: Adverse selection with Pareto distribution of ideas.
$b(\theta) = \frac{3}{(1-\theta)^{\frac{2}{3}}}$; $s(\theta) = \frac{2}{(1-\theta)^{\frac{2}{3}}}$; $\theta \sim U[0, 1]$.

best ideas.

[8]Throughout the analysis, I assume that buyers and sellers engage in a one-shot interaction and cannot contract on ex-post quality. This is consistent with my empirical setting and applies to uncertain, arm's length markets for ideas.

[9]Instead of the competitive equilibrium formulation, one could specify a game-theoretic formulation in which buyers simultaneously offer prices and the seller(s) accepts or rejects. It can be shown that the Pareto-efficient equilibrium described above is the unique SPNE of this game so that buyers cannot get "stuck" at the Pareto-inferior equilibria if the players know $F(\theta)$. Still, inferior equilibria can be relevant if learning about the best equilibria is hard (Scheuer and Smetters [2018]).

When should we expect big markets for ideas to arise? Big markets for ideas go hand in hand with big gains from trade. It is straightforward to show that as the buyer's valuation $b(.)$ rises relatively to the seller's $s(.)$, the proportional and average value of the sold ideas grow.

Proposition 1. *Uniformly higher gains from trade increase the proportion and average value of the ideas traded in equilibrium.*

Proof. In the Appendix. □

However, as well-known in the market for lemons literature, the high-quality goods are harder to sell. The equilibrium price p^*, while appealing to the bad inventors may not satisfy the great ones.

Proposition 2. *[The best ideas remain unsold] Suppose that $\exists\ \hat{\theta}$ s.t. $s(\hat{\theta}) > \mathbb{E}[b(\theta)]$. Then all ideas s.t. $\theta \geq \hat{\theta}$ remain unsold.*

Proof. In the Appendix. □

Arrow's disclosure paradox produces losses $\mathcal{L} \in [0, 1]$ relatively to a perfect information framework because it hinders trade at the top of the distribution. The next result shows that this matters especially when most gains from trade are concentrated there.

For the next result I focus on clockwise rotations of $F(\theta)$ that increase inequality in the value of the ideas by pushing mass from the center towards the tails of $f(\theta)$ (see Fig. 7). As a result, the relative value (and lost gains from trade, assuming a positive relation between the two) of good, unsold ideas grows. Assuming that gains from trade go hand in hand with quality, virtually all value is destroyed ($\mathcal{L} \to 1$) as gains from trade and value increasingly concentrate among few ideas in the right tail.

Proposition 3. *Assume $b' \geq s'$ [10] and equilibrium uniqueness [11]. Inequality-increasing rotations of $F(\theta)$ around $\theta \geq \theta^*$ increase relative losses $\mathcal{L}$.*

Proof. In the Appendix. □

These inequality-increasing rotations effectively increase the relative proportion of *Lemons* vs. *Peaches*. Justin and Myatt [2006] show that any such rotation can be equivalently obtained [12] by adding independent, random noise to θ: rotating F can then be interpreted as making the idea evaluation process "harder" for potential buyers.

[10] This assumption guarantees that inequality-increasing transformations of the idea distribution translate into inequality-increasing transformations of the gains from trade. As a special case, buyers could have a fixed, proportional advantage relative to sellers: $b(\theta) = cs(\theta)$, where $c > 1$.

[11] This boils down to a technical assumption on the log-concavity of the left-hand integral of the cumulative distribution function of the idea distribution. Distributions that satisfy this condition include the normal, the power function, the weibull, the gamma, the pareto and the lognormal, and more in general all distributions with log-concave pdf or cdf (Bagnoli and Bergstrom [2005]).

[12] More precisely, if two distributions F, G share a common mean and are ordered by the (partial) rotation order ($X \prec_R Y$), they are also ordered by second-order stochastic dominance ($X \prec_{SOSD} Y$). Rotschild and Stiglitz [1970] showed that SOSD dominance implies that there must exist a mean independent random noise Z such that $Y = X + Z$.

2.3 NDAs and third-party certification: the role of "disclosure technologies"

Until now, I assumed that sellers have no way to verifiably disclose quality θ. Here I generalize the model by assuming that idea disclosure is possible but "risky" or "costly", for example due to appropriation concerns.

I define a disclosure technology[13] $\mathcal{T} = [\mathbb{L},\ \ell,\ c(\theta)]$ as any device which allows an inventor to disclose verifiable information about his/her idea to the buyer (i) at some cost (ii) with some precision:[14, 15]

- $\mathbb{L} \in [0, 1]$ is the ex-ante probability that an inventor has access to the disclosure technology.
- $\ell \in [0, 1]$ indexes the technology precision according to the monotone information partial order (L. [1988], Athey and Levin [1998]).[16]. Given $\mathcal{T}(\ell'), \mathcal{T}(\ell'')$ with $\ell' > \ell''$, higher signals according to the first technology are interpreted more favorably by the buyer than higher realizations of the second.[17]
- $c(\theta) \geq 0$ is the cost of using the technology. For example, this may represent a fee paid to a third party, the expected loss if the idea is stolen, or the chance that the buyer will refuse to sign the Non Disclosure Agreement.[18]

To clarify the role of each of these parameters, I include in Table A some examples of disclosure technologies.

[13]I do not investigate how a specific set of disclosure technologies arises – some of them arise naturally (e.g: truthful revelation of verifiable information), others may be introduced by a watchdog or the market platform.

[14]Specifically, if the seller θ can and does activate $\mathcal{T}$, the buyer observes a public signal $z \sim G(z|\theta)$ and updates her beliefs about the seller's quality to $F(.|z) \in \Delta(\Theta)$. Signals are stochastically ordered, so that higher z's are good news about quality: $F(\theta|z') \gtrsim_{\text{FOSD}} F(\theta|z)$ for $z' > z$.

[15]Here the disclosure technology refines the buyer's but not the seller's signal. This could be an interesting extension if e.g. inventors apply for a patent to obtain a "second opinion" on the invention value. It is left for future research since it does not mirror the empirical setting which I will analyze.

[16]Given two signals Z, Y which are informative about quality (i.e. $F(.|z), G(.|y)$ satisfying FOSD), $Y \gtrsim_{\text{MIO}} Z$ if for all $q \in [0, 1]$ we have that $G(.|G_Y(Y) \geq q) \gtrsim_{\text{FOSD}} F(.|F_Z(Z \geq q)$.

[17]For example, if Z is obtained simply by adding noise to Y then the latter dominates the former in MIO sense. In the limit, $\ell = 0$ represents a uninformative signal, while $\ell = 1$ represents perfect revelation of θ.

[18]It may also include any pure hedonic value associated with $\mathcal{T}$ (e.g: access to a VC's network; recognition of universal property rights associated with a patent; certification subsidies by the market platform). The assumption on the sign of $c(\theta)$ is driven by the technologies available in the empirical setting to follow.

Table A: Examples of Disclosure Technologies

Technology ($\mathcal{T}$)	Availability ($\mathbb{L}$)	Preciseness (ℓ)	Cost $c(\theta)$
Non-Disclosure Agreement	1	1	$\mathbb{P}_{\text{refuse}}\, b(\theta)$
Third-party revenue verification	$0 < \mathbb{L} < 1$	$0 < \ell < 1$	c
Disclosing proprietary idea	1	1	$\mathbb{P}_{\text{steal}}\, b(\theta)$
Disclosing non-proprietary characteristics	1	1	0
Patents / VC certification	$1_{\theta \geq \theta_o}$	$0 < \ell < 1$	$c(\theta)$

As will be clarified in Section 4, NDAs are available to all inventors, while verification is only available to a subset (hence $0 < \mathbb{L} < 1$).
I assume that NDAs and disclosure perfectly reveal the idea, while revenue verification serves as a noisy signal.
Technologies highlighted in grey are excluded from the empirical analysis.

Equilibrium and unravelling

Only inventors who expect to gain from the disclosure technology $\mathcal{T}$ will use it if available – in the Appendix, I show that under some mild conditions there exists a unique separating perfect bayesian equilibrium, such that every and only inventor above some quality threshold θ_ℓ uses the technology if available (i.e. the "compliers"). Low-quality ideas which dispose of $\mathcal{T}$ (i.e. the "never-takers") and all ideas which are not endowed with $\mathcal{T}$ (the "non-treated") get pooled together and are offered a same price $p_\emptyset$.

Proposition 4. *Assume $b' > c' + s' > 0$. Then $\exists\, \underline{\ell} \in (0, 1]$ s.t. for any sufficiently precise technology ($\ell > \underline{\ell}$) there is a unique, separating perfect bayesian equilibrium: every and only idea of high enough quality $\theta \geq \theta_\ell$ discloses if it is endowed with $\mathcal{T}$.*

Proof. See the appendix. □

Under what conditions does the equilibrium fully unravel (Milgrom [1981]) and everybody who has a disclosure technology will use it? It is straightforward to show that if $\mathbb{L} = 1$ and $c = 0$ then even the worst ideas are forced to disclose and we'll have full unravelling. However, both conditions are necessary: under positive costs or scarce availability of disclosure technology (as in the empirical setting to follow), the worst types do not disclose.

Proposition 5 (Necessary and Sufficient Conditions for Unravelling to Full Disclosure)**.**

$$\mathbb{L} = 1 \textbf{\textit{ AND }} c = 0 \Rightarrow \theta_\ell = \underline{\theta}$$

$$\mathbb{L} \neq 1 \textbf{\textit{ OR }} c \neq 0 \Rightarrow \theta_\ell > \underline{\theta}$$

2.4 The bidding process

In the data, buyers observe and value start-up features beyond just idea quality θ. I generalize the model by including observable characteristics X and structural iid taste shocks (ϵ_b, ϵ_s)$\sim F_\epsilon \perp [\theta, X]$:

$$u_{is}(\theta) = \gamma X + s(\theta) + \epsilon_{is}$$

$$u_{ib}(\theta) = \beta X + b(\theta) + \epsilon_{ib}$$

θ represents idea quality unobserved by buyers: under perfect information, it collapses into a simple constant and is included into X.

Sellers gain access to a disclosure technology with probability $\mathbb{L}(X)$ and conditional on accessing the technology decide whether to pay $c(X, \theta)$ and disclose or not. All potential buyers observe X, their own taste shock ϵ_{ib}, and a noisy signal $\theta_{\mathcal{T}}$ which depends on the disclosure choice. Then they bid[19]

$$p_{ib}(\theta_{\mathcal{T}}) = \beta X + \tilde{b}(\theta_{\mathcal{T}}) + \epsilon_{ib} \tag{5}$$

and the good is sold to the highest bidder at p if $p \geq u_{is}(\theta)$.

I observe the highest bid of the auction p[20], as well as the number of bidders, whether the asset is sold, an ex-post proxy for the quality of the idea (start-up failure), and whether the inventors use a disclosure technology: sellers can ask all potential bidders to sign an NDA ($\mathcal{T}_{\text{NDA}}$), and can verify their (advertising) revenue claims through their Adsense account ($\mathcal{T}_{\text{Verified}}$). Finally, I observe a rich set of covariates $\tilde{X}$ which ought approximate the buyer's information set given by X.[21]

3 Testable implications

The model has clearcut predictions on how disclosure technologies affect speed and safety of the transactions, valuations of both buyers and sellers, and various information-related inefficiencies of a market for business ideas. I explore each of them in the following subsections.

3.1 Disclosure increases market safety

Asymmetric information issues associated with the disclosure problem have the potential to affect the safety of an idea transaction: sellers who refuse to fully disclose an idea have more leeway to miss-portray quality and engage in potentially fraudulent behavior.

Hypothesis 1. *[Less fraudulent behavior among disclosed ideas]*

$$P[\textit{Auction Suspended}|\tau^V] \leq P[\textit{Auction Suspended}|\tau^U]$$

[19] $\tilde{b}$ is increasing in perceived quality $\theta_{\mathcal{T}}$. The specific functional form depends on the auction mechanism.
[20] Given u_{ib}, it is easy to show that p is a linear function of βX and increasing in *perceived* quality.
[21] All expectations and distributions to follow condition on $\tilde{X}$, which I will omit for notational simplicity.

3.2 Disclosed ideas are more valuable and likely to sell

Prop. 4 predicts that not all ideas "treated" with the opportunity to disclose will do so. High-quality ideas benefit from perfect information and will take up the opportunity (the "compliers") whereas low-quality ideas will avoid disclosure (the "never takers") and pool together with the "untreated" ideas. This heterogeneous take-up rate between high and low-quality ideas should be reflected in the buyers' valuations and has stark testable implications.

First, disclosed ideas should be perceived to be of higher-quality by buyers, who will offer higher prices. If availability of the disclosure technologies were randomly assigned, the difference in prices between disclosed and undisclosed ideas would identify the treatment effect of availability among the high-quality compliers.

Hypothesis 2. *[Disclosed ideas command higher bids]* $E(p|\tau^V) \geq E(p|\tau^U)$

At the same time, compliers are also valued more by the sellers: this is purely a selection effect given by the fact that compliers are of higher-quality than untreated and never-takers. In my setting, sellers often set a buy-it-now price at the begin of the auction: if a buyer is willing to offer that amount, the auction ends and the idea is sold immediately. As such, I hypothesize that sellers will require higher buy-it-now prices for verified ideas.

Hypothesis 3. *[Disclosed ideas have higher buy-it-now prices]* $E(\mathit{BINP}|\tau^V) \geq E(\mathit{BINP}|\tau^U)$

At least in principle, gains from trade need not grow with idea quality. If they do ($b' \geq s'$), for instance because bad ideas fail no matter the owner, the price effect will dominate the buy-it-now effect and the distance between the buy-it-now price and the (typically lower) final bid should shrink for disclosed ideas.

Hypothesis 4. *[Sellers of disclosed ideas capture more value]* $E(\mathit{BINP} - p|\tau^V) \geq E(\mathit{BINP} - p|\tau^U)$

For the same reason, if gains from trade are concentrated at the top of the distribution ($b' \geq s'$), disclosed ideas should be sold with higher probability relatively to the pooled group of "untreated" and "never-taker" ideas.

Hypothesis 5. *[Disclosed ideas are more likely to be sold]* $E[1_{sold}|\tau^V] - E[1_{sold}|\tau^U] \geq 0$

3.3 Disclosure assuages information frictions

A popular test for adverse selection in the market for ideas literature (Pisano [1997], Guedj [2005], Danzon et al. [2005], Arora et al. [2009]) is to compare the quality of sold and unsold ideas. The reasoning behind this is that lemons are easier to sell than peaches, so the ex-post quality of sold ideas should be higher than that of unsold ideas. A complication with this test is that it is hard to tell whether ex-post differences in quality between sold and unsold ideas are driven by differences in ex-ante idea quality or differences in how better buyers would manage sold and unsold ideas relatively to sellers. For example, in a market in which the commercialization ability only matters for ideas of (ex-ante) low quality, only low quality ideas would be sold *even under perfect information*: finding a positive gap in the ex-post quality of sold and unsold ideas would therefore be a misleading test for the presence of adverse selection. Viceversa, in a market in which the commercialization ability matters for both good and bad ideas and buyers and sellers transact under imperfect information, only bad ideas are sold: however, the ex-post quality difference between sold and unsold ideas could be similar because only the former are managed by buyers, while the latter are implemented with the sellers' inferior execution strategy. Finding no gap in the ex-post quality of sold and unsold ideas would therefore be a misleading test for the absence of adverse selection.

I generalize this test by looking at differences in quality for sold and unsold ideas *among disclosed and undisclosed ideas*. Under symmetric information, verification should not affect the survival gap between sold and unsold ideas. Under asymmetric information, verification helps to sell the best ideas – so sold, verified ideas are relatively more likely to survive. In the data, I observe how long a start-up survives after the transaction and use this as a measure of ex-post quality.

Hypothesis 6. *[Fewer lemons among disclosed ideas]*

$$E[1_{Death}|\tau^V;1_{Sold}=1]-E[1_{Death}|\tau^V;1_{Sold}=0] \leq E[1_{Death}|\tau^U;1_{Sold}=1]-E[1_{Death}|\tau^U;1_{Sold}=0]$$

To conclude, a last test can be drawn from auction theory and to my knowledge was first suggested within the market for ideas literature by Arora and Gambardella [2010a]. On the platform, multiple buyers bid on a same idea for which they share a common valuation $v(\theta)$ and a private valuation ϵ_i (Eq. 5). Other things equal, the buyer most optimistic about $v(\theta)$ wins the auction: to avoid overpaying (i.e.: the "winner's curse"), bidders shade their bids in proportion to the number of competitors N. Goeree and Offerman [2003] show that public information disclosure (which increases the importance of the private component relatively to the unobserved common component) raises total surplus and the seller's expected revenues. Importantly, the difference in bid elasticity to the number of competitors for private value and common value auctions suggests another test for asymmetric information. When information about the common value is disclosed, the winner's curse bites less: the seller's revenues increase faster as the number of bidders goes up.

Hypothesis 7. *[Disclosure breaks the winner's curse]*

$$\frac{\partial E[p|\tau^V;N]}{\partial N} \geq \frac{\partial E[p|\tau^U;N]}{\partial N}$$

3.4 Endogeneity concerns

In the marketplace, only a proportion $\mathbb{L}(X)$ of lucky sellers can disclose (e.g: only those with a Google Adsense account at the time of the auction). Of these, only the high-quality compliers will choose to disclose. As explained in the section above, I will be then interested in regression equations of the following kind:

$$p_i = \beta \tilde{X}_i + \alpha 1^i_{\mathcal{T}} + e_i \tag{6}$$

where $1^i_{\mathcal{T}}$ denotes disclosure choice: α captures the LATE that the availability of a disclosure technology has on high-quality compliers' price, sale probability, et cetera. $\alpha > 0$ implies that buyers value disclosure and generally transact under imperfect information. This test is only informative, however, if I can dispel concerns of non-excludability and endogeneity of the availability of $\mathcal{T}$ (Assumption 1, Assumption 2).

Assumption 1. *[$\mathcal{T}$ has no intrinsic value under perfect information]*

The excludability requirement is that the disclosure technologies are not valuable per se, independently of whether the buyers are transacting under perfect or imperfect information. A simple placebo test corroborates this: I will show that verification does not affect valuations when buyers bid after signing a Non-Disclosure Agreement.

Placebo Test. *[Verification has no effect on price when buyers ask for NDAs]*

$$E[p|\mathit{NDA}, \mathit{Verified\ Revenues}] = E[p|\mathit{NDA}]$$

Second, there are omitted variable concerns: the choice to disclose $1^i_{\mathcal{T}}(\theta, X)$ is a random variable increasing in quality unobserved by the buyer (which inspires the test $\alpha > 0$ – see Prop. 4), but also influenced by covariates observed by the buyer but potentially unobserved by the econometrician.[22]

Assumption 2. *[No omitted variable bias in disclosure choice]*

$$1^i_{\mathcal{T}}(\theta, X) = 1^i_{\mathcal{T}}(\theta, \tilde{X})$$

This implies that $E[1^i_{\mathcal{T}} e_i | \tilde{X}] = 0$ and allows for a causal interpretation of $\hat{\alpha}$. In words, this requires that disclosure availability and cost do not depend on idea characteristics observable to the buyer but not to the econometrician.

[22] This stems from the fact that both availability $\mathbb{L}(X)$ and disclosure costs $c(\theta, X)$ depend on idea characteristics X such as time, geographical location of buyers and sellers, and industry. For example, businesses which are more likely to digitally record their customer analytics (thus being able to verify their revenues if they choose to do so) might also have invested in a visually more appealing website, in a way observable to potential buyers but not to the econometrician – this may bias $\hat{\alpha}$ upwards.

In practice, I observe the same auction that the potential buyer is bidding on: this allows me to control for a rich set of observables which reduce omitted variable concerns. However, this might not be sufficient.

A complementary approach is to replicate the choice faced by potential buyers in an experimental setting. Instead of exploiting variation in the availability of disclosure, I directly manipulate whether an idea undergoes revenue verification or not through an experiment on Mechanical Turk. In general under symmetric information I would expect to find no effect of verification on elicited prices. I will show that this is not the case: removing third-party verification from an auction listing results in lower valuations.

4 Setting

The ideal empirical setting should feature a market for ideas where:

- Gains from trade exist: buyers and sellers transact over business ideas that are valuable only once properly commercialized. Gains from trade arise because some buyers are better at commercializing the ideas than the original inventors;

- Arrow's disclosure paradox is salient: the potential of the ideas should be hard to assess without the sellers' full collaboration; at the same time, sellers should be concerned that buyers might steal unprotected ideas.

- Pricing data is available: in contrast to the prior literature on the disclosure paradox, this work focuses on estimating welfare losses. This requires a measure of the buyers' and sellers' valuation of the idea (even if the transaction ultimately fails), and how it varies depending on whether disclosure is safe or not.

- All transactions happen publicly and at arm's length: a credible empirical strategy requires the econometrician to have the same (or more) information than the buyers on the idea on sale.

I focus on FLIPPA, a big marketplace which connects buyers and sellers of early-stage online businesses, including apps, SaaS businesses, websites, blogs and eCommerce stores. Given the sheer size of the platform (more than 25k assets sold for over $300m as of 2020), the data provide a unique insight into the consumer and service digital economy (Delgado et al. [2021]).

Gains from trade exist because the businesses are typically sold to entrepreneurs who wish to scale-up the digital infrastructure and grow the idea beyond the concept stage. In results available upon request, I use detailed data on the auctioned websites' technology stack to show that after a successful transaction the acquirers typically streamline the web infrastructure and increase their reliance on cloud service providers. They also expand into new product (e.g.: mobile) and geographic (e.g.: non-U.S.) markets. Finally, they introduce new marketing tools, both to optimize the websites' conversion rate (e.g: from visitor to buyer) and to conduct re-targeting campaigns. All these strategic investments are consistent with a dynamic new ownership focused on scaling-up and monetizing the newly acquired business.

At the same time, the commercial innovations (Png [2017]) underlying an online business are susceptible to expropriation. For example, replicating the digital infrastructure of apps and websites via reverse engineering is a surprisingly easy and widespread practice. Wang et al. [2018] analyzed the metadata of 6m different apps across Google Play and Chinese app stores, and estimated that up to 7% of the apps in their sample are clones created by competitors. Sellers are generally reluctant to disclose trade secrets such as customer lists, business plans and specific financial figures that could be exploited by the competition. The sellers' hesitation to disclose, in turn, complicates an already tough due diligence process faced by potential buyers. The buyers typically deal with first-time sellers and have to complete the transaction without meeting them in person. While the marketplace offers escrow services to protect buyers against scammers, the buyers have typically less than a week to assess the real value of what they bought. Potential acquirers are expressly advised to conduct a thorough due diligence before buying, and to verify that the business does not duplicate content (or otherwise infringe on others' intellectual property) or engage in fraudulent schemes. The decision is often based on information inherently hard to verify[23] or easy to tweak.[24]

The marketplace has introduced two tools to reduce these frictions. First, all sellers may require buyers to sign a Non Disclosure Agreement before accessing key information about the listing (e.g.: the URL or a list of clients). Second, those seller with a Google Adsense account can connect it to the platform so that the marketplace can independently verify their businesses' advertising revenues. Each of these tools represents an archetypal "disclosure technology". NDAs increase protection against expropriation; they are universally available, but costly. Revenue verification provides a noisy signal of the idea's potential; it is free, but only available to eligible businesses.

Finally, an attractive feature of this setting is that most businesses are sold via public auctions. This allows me to collect bidding data on both successful and failed transactions, which I will use to pin down buyers' and sellers' valuations of the idea. Moreover, I have access to the original auction page, which I use to approximate the potential bidders' information set at the time of the sale. To conclude, most online businesses (except those whose URL is confidential) leave a rich digital trace that I can use to track their strategies and outcomes before and after the sale.

I have information revenues, profits, bids, seller characteristics, industry and property type of 14,277 projects on sale between May 2010 and October 2020 (Table B). These include transactions for websites, e–commerce enterprises and other online businesses with at least

[23]E.g: Why is the owner selling? Will the seller help transition the business after the sale?

[24]For example, some sellers might omit labor costs from their profit & loss statement, or inflate income statistics by attributing revenues made by other unrelated businesses.

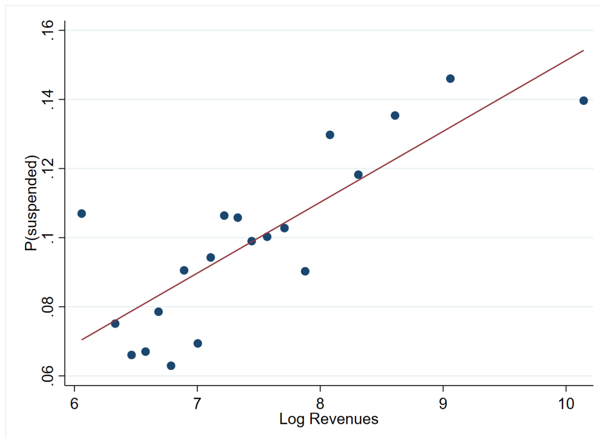
500\$ monthly revenues. Sellers can choose to sell either via an English auction or through private sales; in the first case, they may set a public buy–it–now price and a private reserve price. Around 11% of the listings have verified revenues, while 5% of the websites require bidders to sign an NDA to access further information. I augment the transaction data with information from Builtwith.com on the number and type of technologies employed on the domains associated with the assets on sale: as of November 2020, more than 40% of the businesses has no active technology on the website or returns a "site not found" error.

The high proportion of auctions suspended by the marketplace is a first indication that the market is "unsafe". Figure 2(a) shows that suspension is more likely among auctions claiming to have high revenues[25], suggesting that sellers might misrepresent their quality to persuade potential buyers to bid. In general, selling high-quality ideas on the marketplace is hard. Figure 2(b) shows that high-revenue ideas are less likely to sell. Figure 2(c) highlights that the sellers' valuations (proxied via the buy-it-now price distribution) accelerate faster than the buyers' valuations in the right tail of the revenue distribution. This evokes the "distortion at the top" typical of adversely selected idea markets (Section 2): the next section is devoted to explore this connection.

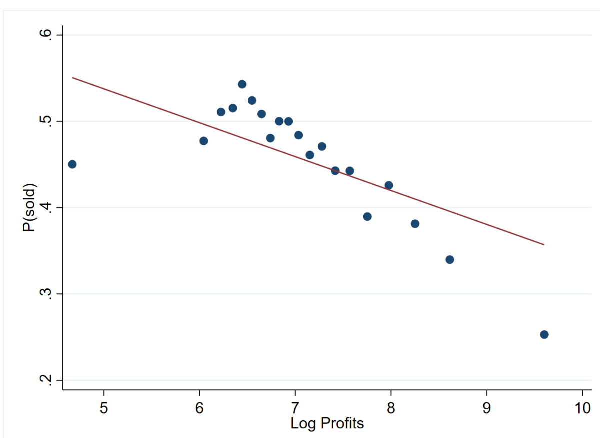
[25]Using profits in lieu of revenues does not affect the results in this section.

Table B: Descriptive Statistics

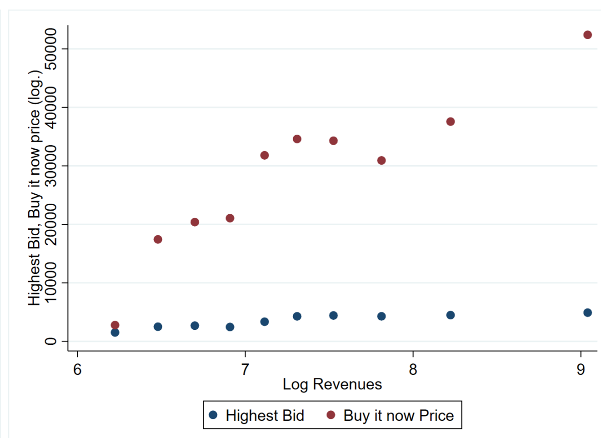
Statistic	N	Definition	Mean	St. Dev.	10^{th} pct.	95^{th} pct.
Sold	14,277	Successful sale	0.443	0.497	0	1
Highest Bid	12,210	Highest bid as of auction closing time	5,853	23,032	0	25,000
Buy It Now Price	10,263	(Optional) desired asking price. Enables immediate purchase at that value	43,789	529671,929	1,050	130,000
NDA Request	14,277	Potential buyers need to sign an NDA to view the asset's URL	0.051	0.221	0	1
Verified Revenues	14,277	Revenues data are verified by Google Adsense	0.109	0.312	0	1
Monthly Profit	14,277	Monthly average profit	6,683	426,167	490	8,755
Monthly Revenues	14,277	Monthly average revenues	8,266	426,549	581	13,805
Bid Count	12,210	Total No. of bids placed	15.5	22.6	0	64
USA	14,277	Seller is based in the United States	0.349	0.477	0	1
Dead (0 elements)	11,425	No live technology found at the asset's web domain on Nov 2020 for Websites founded before Nov 2019 (source: Builtwith.com)	0.418	0.493	0	1
Age at sale (days)	14,277	Days between business establishment and auction ending	804	1,051	79	2,997
Potential	14,277	Seller's pitch contains the keywords *growth*, *investment*, *potential*	0.099	0.299	0	1
Suspended	14,277	Auction suspended by marketplace for reasons including fraud, miss-portrayal of seller identity and assets on sale, refusal to transfer sold assets.	0.099	0.298	0	1



(a) P(auction suspended by the marketplace), by log revenue



(b) P(sale), by log revenue



(c) Highest Bid and Buy-it-now price (med.), by log revenue

Figure 2: Controls include age decile, business model, property type, month-year, repeat sellers and U.S. location fixed effects. Listings with verified revenues and requesting NDA signature are excluded.

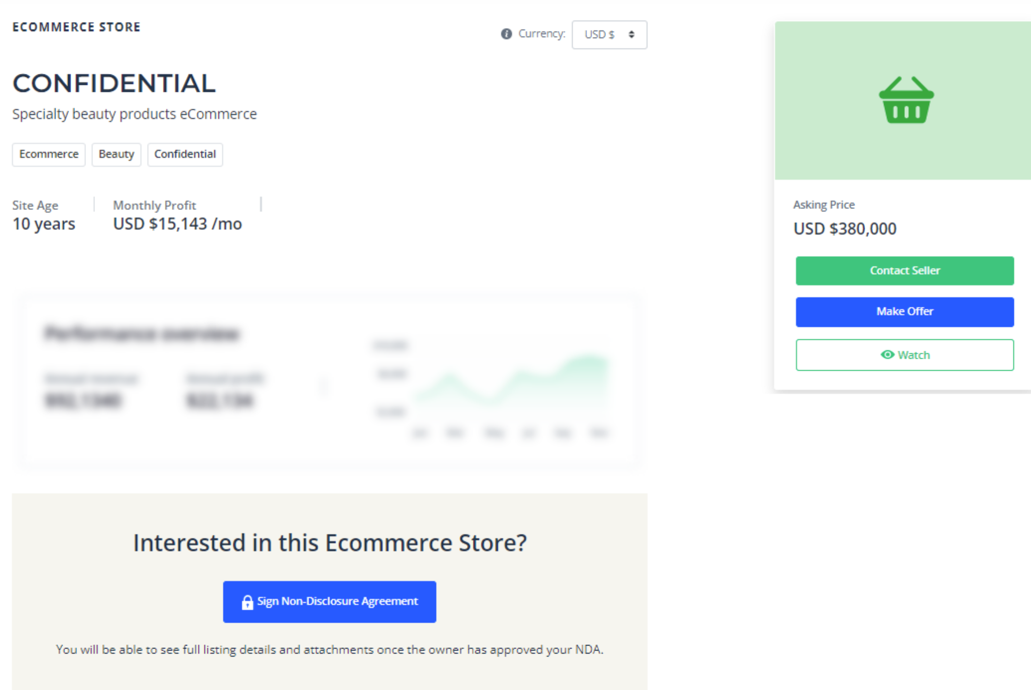


Figure 3: Example of confidential project on sale

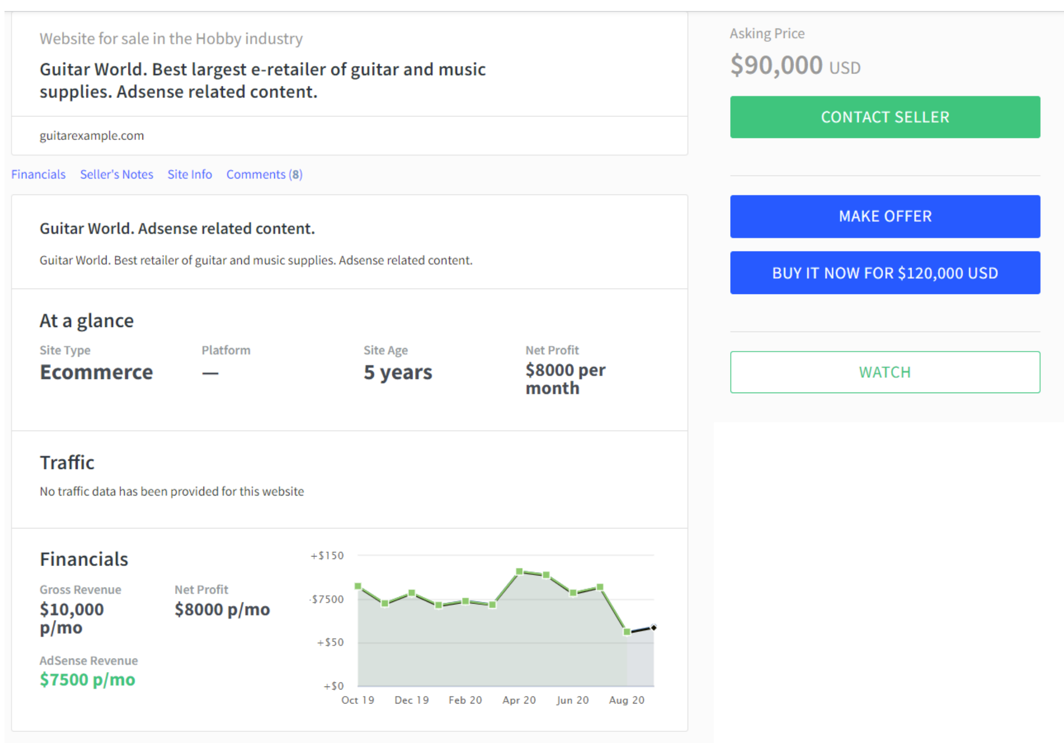


Figure 4: Example of revenue-verified project on sale

5 Results

The main purpose of this section is to test the predictions laid out in Section 3. First, I show that auctions of revenue-verified and NDA-requesting projects are less likely to be suspended by the marketplace compared to other auctions; however, only revenue-verified projects are more likely to sell and elicit higher bids, whereas Non-Disclosure Agreements are not particularly effective in increasing bids and facilitating a sale. In Section 6, I explore this "discouragement" effect in greater detail.

These correlations support the hypothesis that revenue verification (but, perhaps surprisingly, not NDAs) facilitates idea sale by reducing the imbalance of information between buyers and sellers. Additional evidence for this theory is provided by three complementary tests. First, an experiment on MTurk shows that potential buyers bid higher for a verified project compared to an otherwise identical one. Second, I demonstrate that revenue-verified projects alleviate the winner's curse: winners of crowded auctions of revenue-verified ideas are less likely to be concerned about overpaying. A third test shows that sold ideas are more likely to turn out to be lemons, but only if their revenues were unverified at the time of the transaction.[26]

5.1 Disclosure increases the safety of idea transactions

First, I assess whether using disclosure technologies improves market safety (Hypothesis 1). While I do not directly observe whether a seller attempts to defraud a buyer (or viceversa), I can see whether an auction has been suspended by the marketplace.[27] I use this as a proxy for market unsafety, and regress $[1_{\text{Suspended}}]$ on $1_{\text{NDA}}, 1_{\text{Verified}}$ (plus a rich battery of fixed effects and controls). Projects with verified advertising revenues are 12% less likely to have their auctions suspended by the marketplace. Similarly, ideas protected by NDAs are 4% "safer" to sell.

[26] Practical constraints in designing a plausible experiment and following confidential start-ups over time prevent me to replicating two out of the three tests for NDA-requesting projects. The available evidence, however, rejects the notion that NDAs are as effective as revenue verification in assuaging informational frictions.

[27] The reasons behind an auction suspension can be broadly classified in three categories. First, auctions can be suspended if the seller defraudes the buyers (or viceversa) – this includes misrepresenting the business or her identity, shill bidding or failing to deliver the assets (or the payment) after the sale. Second, the auction can be suspended if the seller engages in banned, non-fraudulent behavior (e.g: spam, solicitation or threats). Third, the seller can be suspended if she attempts to defraud the marketplace (e.g: by dealing off-platform or refusing to pay success fees).

Table 1: Disclosure affects Transaction Safety (Hyp. 1)

	P(Suspended)
Verified Revenues	-0.124***
	(0.017)
Confidential	-0.044***
	(0.012)
Constant	0.115***
	(0.002)
Observations	14,381
R-squared	0.203

SE clustered at month, business model level. Also included: decile FE for profit, revenues, age; type of business model and assets on sale; whether the seller is located in the United States and sells for the first time on the platform; monthly FEs.
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

5.2 Verified, but not NDA-protected ideas are more valuable and likely to sell

Second, I look at whether disclosed ideas are valued more by buyers (Hypothesis 2) and sellers (Hypothesis 3). I also look at whether sellers of disclosed ideas capture more value relatively to their own valuations (Hypothesis 4), and whether disclosed ideas are more likely to be sold (Hypothesis 5).

In Column 1 of Table 2 I regress *Log Highest Bid* on *NDA Request* and *Revenue Verification*, plus a rich set of covariates: I find that projects with verified revenues receive substantially higher bids, while projects requesting NDAs command no significant premium.[28] Column 2, however, rejects the notion that revenue-verified projects might be simply more valuable than NDA-requesting ones: sellers request higher buy-it-now prices for verified and, especially, confidential projects. As summarized in Column 3, this implies that sellers of verified projects capture more value (relatively to their own valuations), whereas NDA-requesting sellers are unable to attract enough interest. Table 3 suggests that the effect is driven by those projects which attract no bids: NDAs and revenue verification appear to be equally effective for projects attracting at least one bid.

[28] The average non-verified, non-confidential project receives a $1,800 highest bid.

Table 2: Highest Bid, Buy-it-now Price and Value Capture by Disclosure (Hyp. 2-4)

	(1) Log (High Bid + 1)	(2) Log (Binp + 1)	(3) High Bid / Binp
Verified Revenues	1.168***	0.330***	0.097***
	(0.133)	(0.047)	(0.018)
NDA Request	-0.256	0.780***	0.005
	(0.212)	(0.056)	(0.022)
Constant	5.892***	8.947***	0.427***
	(0.018)	(0.005)	(0.002)
Observations	12,276	10,321	8,944
R-squared	0.073	0.460	0.090

SE clustered at month, business model level. Also included: decile FE for profit, revenues, age; type of business model and assets on sale; whether the seller is located in the United States and sells for the first time on the platform; monthly FEs.
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Table 3: Highest Bid, Buy-it-now Price and Value Capture by Disclosure (Hyp. 2-4) for projects with at least one bid

	(1) Log (High Bid + 1)	(2) Log (Binp + 1)	(3) High Bid / Binp
Verified Revenues	0.746***	0.356***	0.077***
	(0.078)	(0.052)	(0.019)
NDA Request	0.870***	0.778***	0.079**
	(0.106)	(0.064)	(0.024)
Constant	7.425***	8.763***	0.530***
	(0.011)	(0.006)	(0.002)
Observations	9,772	8,614	7,237
R-squared	0.277	0.458	0.082
Bid count	> 0	> 0	> 0

SE clustered at month, business model level. Also included: decile FE for profit, revenues, age; type of business model and assets on sale; whether the seller is located in the United States and sells for the first time on the platform; monthly FEs.
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

In Table 4, I look at whether the higher prices associated with disclosure translate into higher sale probability (Hypothesis 5). This is expected if, for instance, gains from trade are higher among the best ideas ($b' \geq s'$). While revenue-verified businesses are 8% more likely to be sold, this is not the case for projects requiring to sign an NDA.

As discussed more in-depth in Section 6.1, the worse outcomes associated with NDAs are likely caused by a "discouragement" effect, whereby buyers are less likely to bid if required to sign an NDA.

Table 4: Effect of disclosure on the probability to sell (Hyp. 5)

	1_{sold}
Verified Revenues	0.079***
	(0.015)
NDA Request	-0.031
	(0.020)
Constant	0.428***
	(0.042)
Observations	14,382
R-squared	0.170

SE clustered at month, business model level. Always included: decile FE for profit, revenues, age; type of business model and assets on sale; whether the seller is located in the United States and sells for the first time on the platform; monthly FEs.
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

5.3 Proposed explanation: disclosure assuages information frictions

5.3.1 A field experiment on MTurk

As discussed in Section 3.4, the correlations highlighted in the previous section only imply that NDAs and (especially) revenue verification causally affect outcomes (e.g.: highest bids) under Assumption 1 and Assumption 2: the buyers shouldn't value revenue verification or NDA requests *per se* (beyond signalling reasons); the availability and cost of verification and NDAs shouldn't correlate with idea characteristics observed by buyers but not by the econometrician.

I test Assumption 1 via the previously discussed Placebo Test (Table 5): I run a hedonic regression of price on NDA Request, Verified Revenues and their interaction. The interaction between NDA Request and Verified Revenues is strongly negative, and the verification effect drops to zero among confidential projects. This suggests that the two technologies are effectively substitutes, and that they have little standalone value when the buyers are already sufficiently informed. This is unsurprising: while e.g. Venture Capital support provides connections and financial support in addition to reducing informational frictions, NDAs and revenue verification have no clear function when disclosure is safe.

Second, I recreate an experimental setting where Assumption 2 holds for revenue verification.[29] In an experiment conducted on Mechanical Turk, I asked 291 subjects to provide their valuation[30] of a business based on a real auction listing. 128 of them were shown a listing whose revenues had been verified by the marketplace (i.e.: the revenues were displayed in bright green). 163 of them were randomly shown the same listing without verified revenues (i.e.: the revenues were displayed in grey)[31]. As can be seen from Fig 5, the valuations from the certified listing dominate those from the unverified listing. A simple permutation test confirms that the difference in mean valuations between the two groups is statistically significant ($p < 0.05$).

Table 5: Substitutability of Verification and NDAs

	Log High Bid
Verified Revenues	0.751***
	(0.075)
NDA Request	0.983***
Verified#NDA	-0.536^{+}
Constant	5.183***
	(0.222)
Observations	10,083
R-squared	0.292

SE clustered at month, business model level. Also included: decile FE for profit, revenues, age; type of business model and assets on sale; whether the seller is located in the United States and sells for the first time on the platform; monthly FEs.
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

[29]Cost is zero both on the platform and in the experiment, hence trivially orthogonal to unobservables. In the experiment, however, I can randomly assign the possibility to verify to two otherwise identical listings – hence availability is orthogonal as well.

[30]Specifically, they were presented with five options: $\$0 - 500$; $\$500 - 5k$; $\$5k - 10k$; $\$10k - 20k$; $> \$20k$.

[31]In the experiment, I included the same explanation of the difference between green and grey figures that is usually featured in the auction text. An attention check at the end of experiment showed that most respondents understood this feature, and the few who did not were removed from the experiment.

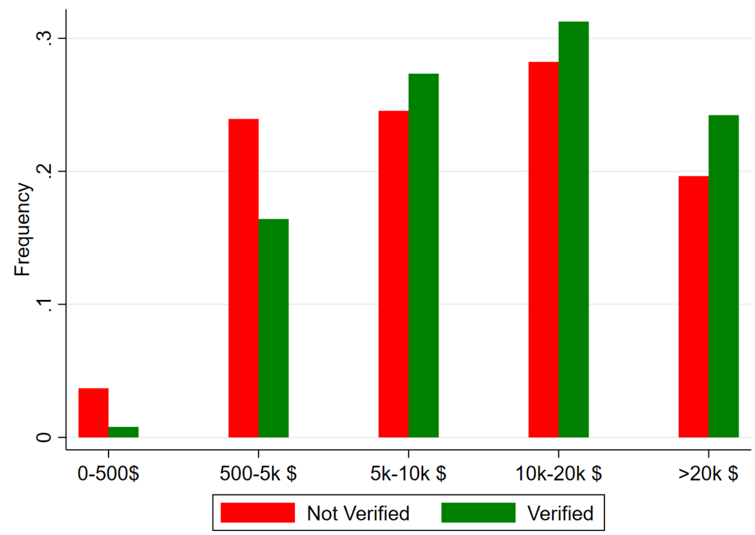


Figure 5: Distribution of valuations by MTurk respondents (Likert scale).

5.3.2 Unverified, but not verified ideas are disproportionately lemons

In this section I investigate differences in ex-post quality of sold and unsold projects, depending on whether they were verified or not during the auction. The rationale is that unverified ideas are subject to a market for lemons issue: the projects sold should turn out to be disproportionately lemons due to adverse selection; asymmetric information, however, is less pronounced if the ideas were verified during the transaction: lemons should be less typical among sold, verified ideas than in the sold, unverified population.

For non-confidential projects I can observe the website's URL and track whether the website later shuts down[32]: I use this to proxy for business failure and distinguish between projects which proved ex-post to be peaches from those which proved to be lemons.[33] I test whether verified ideas are less likely to turn out to be lemons (Hypothesis 6) by regressing 1_{death} on 1_{sold}, 1_{verified}, and their interaction. I include the usual batter of fixed effects and project-level covariates (including month Fixed Effects). The results confirm that verification assuages the lemons issue: sold projects are more likely to fail than unsold ones, but the relationship is reversed among verified ideas.[34]

[32]Unfortunately, the URL is not available for confidential assets.

[33]In practice, I rely on Builtwith.com data via the website's URL and classify a website as dead if no live technologies were found on the website by November 2020. This allows me to run a few robustness checks to make sure, e.g., that I'm not including websites still alive but likely resold for scrap (by tweaking the threshold number of technologies used to classify a website as alive and functioning, or not), or that the results are robust to different definitions of website death (e.g. within 6 months or 1 year).

[34]In principle, acquirers might kill the projects (Cunningham et al. [2021]) or repurpose the underlying idea even if the website dies. While I cannot definitely rule this out, I do not find evidence that buyers of lemon projects abruptly stop investing in the digital infrastructure of the website soon after the auction; moreover, it is unclear why this ought happen disproportionately more often to non-disclosed projects.

Table 6: Disclosed ideas are less likely to be lemons (Hyp. 6)

	P(Death)
Verified Revenues	-0.089***
	(0.015)
1_{sold}	0.045***
	(0.010)
Verified#1_{sold}	-0.049*
	(0.023)
Constant	0.431***
	(0.066)
Observations	11,425
R-squared	0.182

SE clustered at month, business model level. Also included: decile FE for profit, revenues, age; type of business model and assets on sale; whether the seller is located in the United States and sells for the first time on the platform; monthly FEs.
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

5.3.3 Disclosure decreases the fear to overpay

Table 7 summarizes the results about the winner's curse (Hypothesis 7). For projects with verified revenues, there's strong evidence that bids raise faster with the number of bidders relatively to unverified projects – again, this can be interpreted as evidence that verification is informative, and turns a common auction (where bidders are afraid to "overpay") into a private one (where bidders know what they are bidding for). Evidence is not as strong for confidential projects.

Table 7: Disclosure breaks the winner's curse (Hyp. 7)

	Log High Bid
Log Bid Count	0.215***
	(0.020)
Verified Revenues	0.260
	(0.166)
Verified#Log Bid Count	0.182***
	(0.051)
NDA Request	0.813***
	(0.231)
NDA#Log Bid Count	0.053
	(0.081)
Constant	4.738***
	(0.232)
Observations	9,824
R-squared	0.303

SE clustered at month, business model level. Also included: decile FE for profit, revenues, age; type of business model and assets on sale; whether the seller is located in the United States and sells for the first time on the platform; monthly FEs.
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

6 Cost and availability of disclosure technologies

6.1 Few sellers use the disclosure technologies. Why?

The evidence provided above suggests that revenue verification (especially) and non-disclosure agreements (to a lesser extent) help assuage information frictions. Why, then, are only 5% of the start-up ideas protected by NDAs and only 11% of the projects have verified revenues?

According to Proposition 5, a disclosure technology is widely adopted only when two necessary conditions are met. First, it needs to be inexpensive. Second, it needs to be universally available.

Revenue verification, on the one hand, is free but only available to businesses with advertising revenues. This imperfect availability prevents non-advertising businesses to signal their quality; moreover, it allows subpar advertising businesses to avoid verification and hide their low quality to potential bidders. To investigate this, I exploit Builtwith.com data to identify businesses which have been digitally recording their sales and customer analytics before the auction.[35] In other words, I have a proxy for whether a business was "treated" with the opportunity to disclose. I can then directly test whether treated businesses which decided not to disclose (plausibly unbeknownst to buyers) are disproportionately of lower quality: are "non-compliers"

[35]Builtwith.com records the type of technologies hosted on a website hosts. This information is not reported during the auction and is plausibly unobserved by potential buyers at the time of the transaction.

more likely to be lemons and fail after the transaction?[36] The data suggest that this is indeed the case (Table 8): unverified ideas which could have been verified are 20% more likely to turn out to be lemons. I explore how a more widespread availability of revenue verification affects welfare in Section 6.2.

Table 8: Non-compliers are more likely to be lemons

	P(Death)	P(Death)	P(Death)
Verified#$1_{\text{can verify}}$	-0.195***	-0.211***	-0.402*
	(0.033)	(0.052)	(0.179)
Verified Revenues	0.027	0.010	0.245
	(0.030)	(0.046)	(0.172)
$1_{\text{can verify}}$	0.419***	0.492***	0.511***
	(0.013)	(0.017)	(0.022)
Constant	0.147	-0.116	-0.158
	(0.093)	(0.112)	(0.120)
Observations	11,425	5,157	3,132
R-squared	0.294	0.325	0.343
Only Sold	NO	YES	YES
Only non-advertising	NO	NO	YES

SE clustered at month, business model level. Also included: decile FE for profit, revenues, age; type of business model and assets on sale; whether the seller is located in the United States and sells for the first time on the platform; monthly FEs.
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Availability, on the other hand, is unlikely to be the reason NDAs are scarcely used: the marketplace offers all sellers the possibility to request buyers to sign non-disclosure agreements. Table 9 offers evidence that NDAs impose significant, indirect costs on sellers: ideas protected through NDAs receive fewer bids and are more likely to have no bidder at all. In other words, there seems to be a "discouragement effect" which deters from bidding and disincentivizes sellers from asking buyers to sign NDAs (more on this in Section 6.3). While only suggestive, this is consistent with anecdotal accounts of Venture Capital investors refusing to sign NDAs before early-stage pitches .[37]

[36] I define $1_{\text{can verify}}$ to be 1 if the website had ads technologies before the auction. Among the 11,425 businesses in the baseline specification, $P(\text{Verified}|1_{\text{can verify}} = 1) = 12\%$ and $P(\text{Verified}|1_{\text{can verify}} = 0) = 6\%$. This suggests that ads technologies are a strong yet imperfect proxy for the possibility to verify revenues.

[37] https://www.forbes.com/sites/ericwagner/2013/08/27/entrepreneur-vs-venture-capitalist-the-curse-of-the-nda/

Table 9: Projects requesting NDA signature have fewer bids

	(1) Bid Count	(2) No Bids
Verified Revenues	3.406***	-.070***
	(0.961)	(0.014)
NDA Request	-3.796***	0.100***
	(1.053)	(0.022)
Constant	19.151***	0.207***
	(2.958)	(0.002)
Observations	12,277	12,277
R-squared	0.121	0.117

SE clustered at month, business model level. Always included: decile FE for profit, revenues, age; type of business model and assets on sale; whether the seller is located in the United States and sells for the first time on the platform; monthly FEs.
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

6.2 Lost gains from trade in the absence of revenue verification

In this section I estimate lost gains from trade when revenue verification is shut down. My strategy is to compare the demand and supply of start-up ideas for verified and unverified ideas while taking into account that only high-quality ideas are verified in equilibrium.

I assume that the sellers' and buyers' idea valuations can be separated into[38] idiosyncratic taste shocks (ϵ_s, ϵ_b drawn from $F_\epsilon \perp \theta$), and average valuations $s(\theta)$ and $b(\theta)$. The key quantities that I need to identify are $s(\theta), b(\theta), F_\theta$ and F_ϵ. Briefly, when an idea is unverified, the buyers' bid distribution reflects variation in ϵ_b which identifies F_ϵ; when an idea is verified, quality becomes observable and the bid distribution reflects variation in $\epsilon_b + b(\theta)$: my strategy is to back out the distribution of $b(\theta)$ by deconvoluting the latter distribution from the former. $F_{s(\theta)+\epsilon_s}$ can be pinned down from the sale probability of unverified ideas: $F_{s(\theta)+\epsilon_s}(x) = \mathbb{P}(\text{Sale at price } x|\text{Unverified})$.[39] $F_{s(\theta)}$ obtains from $F_{s(\theta)+\epsilon_s}$ via deconvolution and knowledge of F_ϵ. In the appendix I show that this is sufficient to identify $s(\theta)$, $b(\theta)$ and $F(\theta)$ under the following assumptions:

1. the availability (but not use) of revenue verification is independent of quality θ.

2. revenue verification perfectly reveals the idea: under this assumption, $b(\theta)$ can be identified from the bid distribution of verified ideas; relaxing this conservative assumption would increase the welfare losses in the imperfect information scenario relatively to the perfect information one.

[38]Idea valuations also depend on observable idea characteristics X. I ignore this in the exposition and always flexibly control for X in the estimation.

[39]Here I use $\epsilon_b \perp [\theta, \epsilon_s]$.

3. $b(\theta)$ is proportional to $s(\theta)$. While in principle I could estimate a non-parametric lower bound for lost gains from trade, this assumption greatly aides in the estimation.

4. $\epsilon_s \perp \epsilon_b$: conditional on observables, demand and supply can be traced without the aid of instrumental variables. Relaxing this assumption would require to find instruments producing enough variation in prices to trace supply and demand along the whole distribution of θ.

Ultimately, these assumptions allow me to understand how switching verification on and off (that is, enabling buyers to observe θ) affects gains from trade.

Figure 6 plots the empirical distribution of the (residualized) buyers' valuation $b(\theta)$. The distribution is very skewed and unequal: I estimate a Gini coefficient of 0.81. As a benchmark, Gambardella et al. [2008] examine a sample of 8,217 patents from the PatVal-EU survey: they find the value distribution of the patents to be extremely skewed, consistent with Log-Normal distribution characterized by a Gini coefficient of 0.84.

All in all, the calibration exercise suggests that shutting down revenue verification destroys 63.6% of potential gains from trade, reduces welfare[40] by 14.7% and leads to 17.7% fewer ideas to be traded[41]. A first point worth mentioning is that not all gains from trade are lost when disclosure is shut down: the market does not unravel completely because the buyer values the bad ideas enough more than the seller. Second, the lost gains from trade (-64%) are disproportionately big relative to the decrease in the number of traded ideas (-18%): this happens because the lost ideas are disproportionately high-value, reflecting the fact that the best ideas are the hardest to sell under imperfect information.

To further stress this point, I simulate lost gains from trade under three counterfactual distributions of $b(\theta)$ and $s(\theta)$ (Table 10): idea quality is drawn from a Log-Normal with the same mean but different Gini inequality. When the inequality is high (Gini index: 0.95) the lost gains from trade under no verification amount to 65%, while the number of traded ideas decreases by only 8%. With average inequality (Gini index: 0.5), the lost gains from trade are around 40% while there is a decrease of 17% in the number of traded ideas. Finally, when the inequality is low (Gini index: 0.25) less than 10% of gains from trade are lost and I find a 5% decrease in the number of ideas sold relatively to the perfect information benchmark. Again, this shows that high inequality exacerbates the losses even if only few ideas are affected by the disclosure paradox.

[40] Welfare losses are lower because they include in the denominator the hedonic value of observable idea characteristics X.

[41] The 95% bootstrapped confidence interval of lost gains from trade is $[61.2\%, 65.9\%]$; the 95% bootstrapped CI of lost welfare is $[14.0\%, 15.4\%]$; the 95% bootstrapped CI of fewer sold ideas is $[15.6\%, 19.0\%]$.

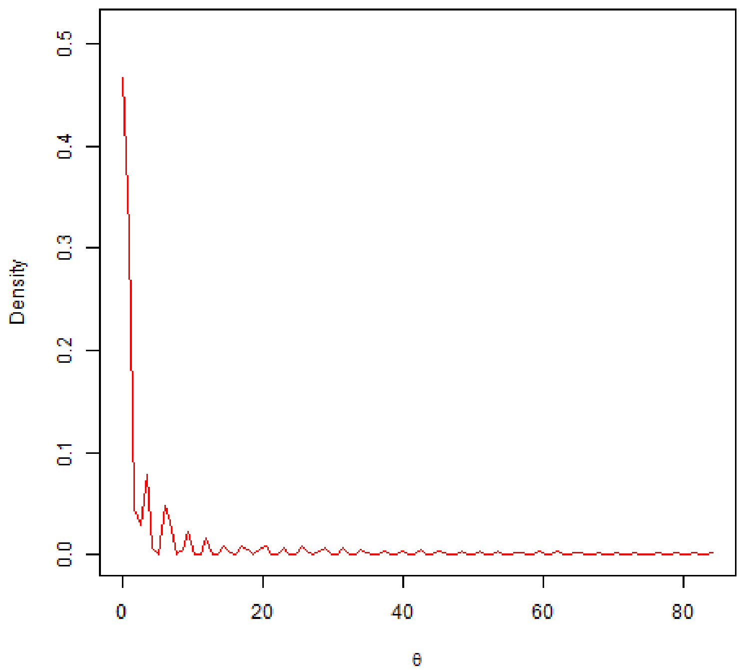


Figure 6: Empirical distribution of idea quality θ

Table 10: The inequality of the idea distribution affects welfare losses

	Gini inequality		
	0.95	**0.50**	**0.25**
Gains from trade	-64.8%	-40.1%	-9.9%
Sold ideas	-7.6%	-16.7%	-4.4%

Here θ is drawn from three Log-normal distributions with the same mean as $\hat{F}(\theta)$, but different Gini inequality. Losses are relatively to a perfect information benchmark.

6.3 How likely are potential buyers to sign non-disclosure agreements?

Table 9 shows that confidential projects are 1.5x less likely than non-confidential projects and 2.1x less likely than projects with verified revenues to receive at least one bid. This suggests that buyers dislike NDAs and are less likely to bid when they are required to sign one. These figures, however, severely underestimate the cost that NDAs impose on sellers if, for example, gains from trade grow with quality or if the marginal value of bids is higher for high-quality ideas[42]. To make progress, I take a complementary approach and estimate what NDA cost would rationalize the limited use of NDAs I witness in the data.

I assume that using NDAs involves a fixed cost d and prevents a sale with probability c.[43] From Prop 4, only a subset of the best inventors will use the NDAs, whereas the others will refuse to disclose and transact under asymmetric information. Labelling θ^* as the quality threshold above which NDAs are requested, the equilibrium price in the lower market is $p = E[b(\theta)|b(\theta)+\epsilon_b \geq s(\theta)+\epsilon_s; \theta \leq \theta^*]$. At the threshold quality, an inventor is indifferent between disclosing or not: $p = (1-c)b(\theta^*) - d$. Knowledge of $b(\theta), s(\theta), \epsilon$ comes from the previous section. Moreover, in the data I observe that 5% of the inventors use NDAs, so I set θ^* s.t. $b(\theta^*)$ is at the 95% quantile of the $b(\theta)$ distribution. This allows me to calculate p, and estimate c under an assumption on d.

Intuitively, the fact that only 5% of the inventors decide to disclose is consistent with sizable NDA costs: all sellers would use free NDAs. Assuming $d = 0$, I estimate $c = 97.6\%$ (95% bootstrapped CI: $[97.4\%, 97.9\%]$). Assuming fixed costs equal to half the mean quality θ, $c = 89.3\%$ (95% bootstrapped CI: $[88.7\%, 90.0\%]$) – this amounts to a reduction in the probability of selling the idea by 89%[44].

[42]In the first case, quality would be positively correlated with both NDA use and the number of bids – resulting in a downward estimate of the negative effect of NDAs on the number of bids. In the second case, a 1.5x decrease in the number of bids translates in an even lower likelihood of sale for high-quality ideas.

[43]Specifically, with probability c the sellers receive bids lower than their reservation wage, whereas with probability $1-c$ they are able to attract a buyer willing to pay full value $b(\theta)$.

[44]Holding idea characteristics, valuations and information fixed.

7 Conclusions

During the COVID-19 crisis, a timely agreement between BioNTech and Pfizer allowed for a more rapid commercialization of the mRNA vaccines: Schneider et al. [2021] estimate that a one-year delay in the U.S. vaccination campaign would have resulted in about 1.1 million additional deaths. This is but one example of why efficient markets for technologies are so important: they allocate great ideas to those best positioned to commercialize them, thus maximizing the impact of innovations.

In this work I focused on how markets for ideas can unravel if the inventors don't disclose their ideas to potential buyers because of appropriation concerns. I showed that this results in adverse selection: buyers do not buy the high-quality ideas because they are unable to recognize them. Then, I highlighted that disclosure issues cause huge welfare losses when most gains from trade are concentrated among the best ideas: this happens when good ideas are rare and very valuable when commercialized in the right way, while bad ideas fail independently from the owner.

I then showed that this issue is salient in a marketplace of early-stage online businesses, providing evidence that (especially) revenue verification facilitates trade and increases inventors' returns. I find that bad sellers behave strategically, avoiding disclosure to hide behind a veil of uncertainty. This is detrimental for welfare: a calibration of the model to the data reveals that shutting down disclosure through revenue verification destroys 63% of the potential gains from trade, concentrated among a handful of crucial, failed transactions. Non-disclosure agreements are less useful than expected, because they discourage many potential buyers from bidding.

Beyond static losses, the disclosure paradox reduces the incentives to innovate, and more specifically the returns from producing breakthrough ideas: it's plausible that an efficient market for ideas would increase both the number and the average quality of innovations, positively impacting economic growth. This seems an interesting avenue for future work.

References

Akerlof, G. A. (1970). **The Market for "Lemons": Quality Uncertainty and the Market Mechanism**. *The Quarterly Journal of Economics*, 84(3):488–500.

Anton, J. J. and Yao, D. A. (1994). **Expropriation and Inventions: Appropriable Rents in the Absence of Property Rights**. *American Economic Review*, 84:190–220.

Arora, A. and Gambardella, A. (2010a). **Ideas for rent: an overview of markets for technology**. *Industrial and Corporate Change*, 19(3):775–803.

Arora, A. and Gambardella, A. (2010b). **The Market for Technology**. *Handbook of the Economics of Innovation, in: Bronwyn H. Hall & Nathan Rosenberg (ed.), Handbook of the Economics of Innovation*, 1(1):641–678.

Arora, A., Gambardella, A., Magazzini, L., and Pammoli, F. (2009). **A breadth of fresh air? Firm type, scale, scope and selection effects in drug development**. *Management Science*, 55(10):1638–1653.

Arrow, K. (1962). **Economic Welfare and the Allocation of Resources for Invention**. *Rate and Direction of Inventive Activity: Economic and Social Factors*, 1:609–626.

Athey, S. and Levin, J. (1998). **The Value of Information in Monotone Decision Problems.** Working Paper, Department of Economics MIT.

Azevedo, E. M., Deng, A., Olea, J. M., Rao, J., and Weyl, E. (2020). **A/B Testing with Fat Tails**. *Journal of Political Economy*, 128(12):4614–000.

Bagnoli, M. and Bergstrom, T. (2005). **Log-concave probability and its applications**. *Economic Theory*, 26:445–469.

Biais, B. and Perotti, E. (2008). **Entrepreneurs and new ideas**. *RAND Journal of Economics*, 39(4):1105–1125.

Cunningham, C., Ederer, F., and Ma, S. (2021). **Killer Acquisitions**. *Journal of Political Economy*, 129(3):649–702.

Danzon, P. M., Nicholson, S., and Pereira, N. (2005). **Productivity in pharma-biotechnology R&D: The role of experience and alliances**. *Journal of Health Economics*, 24(2):317–339.

Delgado, M., Kim, J. D., and Mills, K. (2021). **The Servicification of the U.S. Economy: The Role of Startups versus Incumbent Firms**. NBER Working Paper.

Gambardella, A., Harhoff, D., and Verspagen, B. (2008). **The value of European patents.** *European Management Review*, 5(2):69–84.

Gans, J., Hsu, D. H., and Stern, S. (2008). **The Impact of Uncertain Intellectual Property Rights on the Market for Ideas: Evidence from Patent Grant Delays**. *Management Science*, 54(5):982–997.

Gans, J. S. and Stern, S. (2003). **The product market and the market for "ideas": commercialization strategies for technology entrepreneurs**. *Research Policy*, 32:333–350.

Giuri, P., Mariani, M., Brusoni, S., Crespi, G., Francoz, D., Gambardella, A., Garcia-Fontes, W., Geuna, A., Gonzales, R., Harhoff, D., Hoisl, K., Bash, C. L., Luzzi, A., Magazzini, L., Nesta, L., Nomaler, O., Palomeras, N., Patel, P., Romanelli, M., and Verspagen, B. (2007). **Inventors and invention processes in Europe:Results from the PatVal-EU survey**. *Research Policy*, 36:1107–1127.

Goeree, J. K. and Offerman, T. (2003). **Competitive Bidding in Auctions with Private and Common Values**. *The Economic Journal*, 113(489):598–613.

Grossman, S. J. (1981). **The Informational Role of Warranties and Private Disclosure about Product Quality**. *Journal of Law and Economics*, 24(3):461–483.

Guedj, I. (2005). **Ownership vs. Contract: How Vertical Integration Affects Investment Decisions in Pharmaceutical R&D**. University of Texas Austin Business School Working Paper.

Harhoff, D., Narin, F., Scherer, F., and Vopel, K. (1999). **Citation Frequency and the Value of Patented Inventions**. *The Review of Economics and Statistics*, 81(3):511–515.

Hegde, D. (2014). **Tacit knowledge and the structure of license contracts: Evidence from the biomedical industry**. *Journal Econom. Management Strategy*, 23(3):568–600.

Hegde, D. and Luo, H. (2017). **Patent Publication and the Market for Ideas**. *Management Science*, 64(2):652–672.

Hsu, D. H. (2006). **Venture Capitalists and Cooperative Start-up Commercialization Strategy**. *Management Science*, 52(2):204–219.

Ibragimov, I. A. (1956). **On the composition of unimodal distributions**. *Teor. Veroyatnost. i Primenen.*, 1:283–288.

Jovanovic, B. (1982). **Truthful Disclosure of Information**. *Bell Journal of Economics*, 13(1):36–44.

Justin, J. P. and Myatt, D. P. (2006). **On the Simple Economics of Advertising, Marketing, and Product Design.** *American Economic Review*, 96(3):756–784.

L., L. E. (1988). **Comparing Location experiments**. *Ann. Statist.*, 16(2):521–533.

Levin, J. (2001). **Information and the market for lemons**. *RAND Journal of Economics*, 32(4):657–666.

Lewis, G. (2011). **Asymmetric Information, Adverse Selection and Online Disclosure: The Case of eBay Motors**. *American Economic Review*, 101:1535–1546.

Luo, H. (2014). **When to sell your idea: Theory and evidence from the movie industry**. *Management Science*, 60(12):3067–3086.

Mas-Colell, A., Whinston, M., and Green, J. (1995). ***Microeconomic Theory***. New York: Oxford University Press.

Megginson, W. L. and Weiss, K. A. (1991). **Venture Capitalist Certification in Initial Public Offerings**. *The Journal of Finance*, 46(3):879–903.

Milgrom, P. and Weber, R. (1982). **A Theory of Auctions and Competitive Bidding**. *Econometrica*, 50(5):1089–1122.

Milgrom, P. R. (1981). **Good News and Bad News: Representation Theorems and Applications**. *Bell Journal of Economics*, 12(2):380–391.

Nelson, R. (1959). **The Simple Economics of Basic Scientific Research**. *Journal of Political Economy*, 67:297–306.

Pakes, A. (1986). **Patents as Options: Some Estimates of the Value of Holding European Patent Stocks**. *Econometrica*, 54:755–784.

Pisano, G. (1997). **R&D Performance, Collaborative Arrangements, and the Market for Know-How: A Test of the "Lemons" Hypothesis in Biotechnology**. Harvard Business School Working Paper.

Png, I. P. L. (2017). **Law and Innovation: Evidence from State Trade Secrets Laws**. *The Review of Economics and Statistics*, 99(1):167–179.

Rotschild, M. and Stiglitz, J. (1970). **Incresing Risk: I. A Definition**. *Journal of Economic Theory*, (2):225–243.

Scheuer, F. and Smetters, K. (2018). **How Initial Conditions Can Have Permanent Effects: The Case of the Affordable Care Act**. *American Economic Journal: Economic Policy*, 10(4):302–343.

Schneider, E. C., Shah, A., Sah, P., Moghadas, S. M., Vilches, T., and Galvani, A. P. (2021). **The U.S. COVID-19 Vaccination Program at One Year: How Many Deaths and Hospitalizations Were Averted?**

Teece, D. (1986). **Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy**. *Research Policy*, 15:285–305.

Wang, H., Liu, Z., Liang, J., Vallina-Rodrigues, N., Guo, Y., Li, L., Tapiador, J., Cao, J., and Xu, G. (2018). **Beyond Google Play: A Large-Scale Comparative Study ofChinese Android App Markets**. 2018 Internet Measurement Conference, https://arxiv.org/pdf/1810.07780.pdf.

Åstebro, T. and Serrano, C. (2015). **Business Partners: Complementary Assets, Financing, and Invention Commercialization**. *Journal of Economics Management Strategy*, 24(2):228–252.

1

Appendix A: Omitted Proofs

Proposition 1: Gains from trade and the market for ideas

Take a strictly positive monotone transformation $\tilde{v}(\theta)$ of $v(\theta)$. Then, $\tilde{\theta}^* > \theta^*$, since:

$$\mathbb{E}[v(\theta)|r(\theta) \leq p] < \mathbb{E}[\tilde{v}(\theta)|r(\theta) \leq p] \quad \forall p > r(\underline{\theta})$$

This implies that both the proportion of ideas sold $F(\theta^*)$ and the average value to sellers $\mathbb{E}[r(\theta)|\theta < \theta^*]$ and buyers $\mathbb{E}[v(\theta)|\theta < \theta^*]$ increase.

Proposition 2: The best ideas remain unsold

Given an equilibrium price p^* and $r(\theta)$ weakly increasing, the statement trivially follows from:

$$p^* = \mathbb{E}[v(\theta)|r(\theta) \leq p^*] \leq \mathbb{E}[v(\theta)] < r(\hat{\theta})$$

Proposition 3: Increases in inequality of the idea distribution

I focus on an ordered family of cumulative distribution functions $F_s(\theta)$ indexed by s. Following Justin and Myatt [2006], I define a (clockwise) rotation as follows:

Definition. *A local change in s leads to a rotation of $F_s(\theta)$ if, for some θ_s^* and each $\theta \in (\underline{\theta}_s, \overline{\theta}_s)$*

$$\theta \lesseqgtr \theta_s^* \Leftrightarrow \frac{\partial F_s(\theta) \Leftrightarrow}{\partial s} \lesseqgtr 0$$

If this holds for all s, then $\{F_s(\theta)\}$ is ordered by a sequence of rotations.

While rotations represent a more flexible notion of increased dispersion than a spread (Rotschild and Stiglitz [1970]), the two order relations are tightly connected: Justin and Myatt [2006] show that if $\{F_s(\theta)\}$ share a common mean then the rotation relative order implies SOSD; on the other hand, if a finite sequence of distributions is ordered by SOSD, then each can be obtained from any other by adding a sequence of mean-preserving spreads (i.e. "simple" rotations).

I additionally impose two assumptions on the rotation-ordered $\{F_s(\theta)\}$ family I will focus on. First, I require all of them to share the same mean, to underscore the notion that rotations affect the preciseness of the signal of the buyer, not the average quality of the idea. Second, I require all of them to be log-concave: a sufficient condition for this is that $F_{s'}$ can be obtained by F_s by adding log-concave random noise if $s' > s$ (Ibragimov [1956]). This guarantees that rotations preserve the equilibrium uniqueness of the model presented in the paper.

To prove Prop. 3, all I need to show is that any mean-preserving rotation of F around θ_s^* necessarily lowers $E[\theta|\theta \leq t]$ for any $t \leq \theta_s^*$, hence losses relatively to perfect information

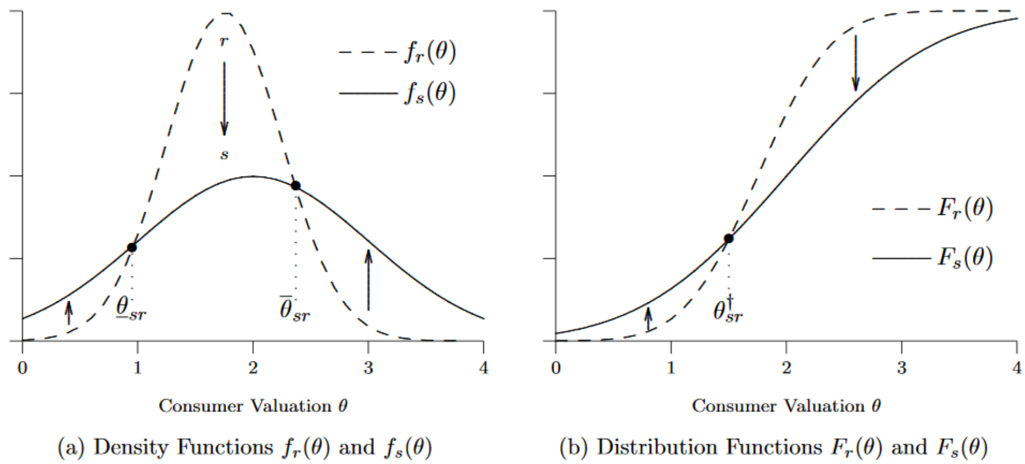


Figure 7: Spread and Rotation for the Normal Distribution.
Reproduced from Justin and Myatt [2006].

($\mathcal{L} = \frac{E[b-r(b)|b\geq b^*](1-F(b^*))}{E[b]}$) increase. Define F_b as the result of the mean-preserving rotation of F_a. Then F_a SOSD dominates F_b, which implies that $\forall t$:

$$\int_{-\infty}^{t} F_b(x)dx \geq \int_{-\infty}^{t} F_a(x)dx \quad \Leftrightarrow \quad E_b[x|x \leq t]F_b(t) \leq E_a[x|x \leq t]F_a(t)$$

Finally, $F_a(t) \leq F_b(t) \Leftrightarrow t \leq \theta^*_s$. This implies that $E_b[x|x \leq t] \leq E_a[x|x \leq t]$ if $t \leq \theta^*_s$.

Proposition 4 (Unique separating Perfect Bayesian Equilibrium)

I assume that $b' > c' + s' > 0$ and that sellers are free to dispose of $\mathcal{T}$ if available.

If a seller of quality θ is endowed with $\mathcal{T}$ and decides to use it, she draws a random signal $x \sim G(x|\theta)$ with support X. By assumption, $G(x|\theta')$ first order stochastically dominates $G(x|\theta)$ whenever $\theta' > \theta$. When the buyers observe higher $x's$, they bayesianly update their belief about θ:

$$\tilde{f}(\theta|x) = \frac{g(x|\theta)f(\theta)}{\int_{\Theta^*} g(x|\theta)f(\theta)d\theta} \tag{7}$$

where Θ^* is the set of types which use the disclosure technology in equilibrium. Because of FOSD, posteriors which condition on higher $x's$ dominate posteriors conditioning on lower $x's$. Viceversa, if the seller fails to draw $\mathcal{T}$ or secretly discards it, the buyers observe no signal and update accordingly. Proving that there is a unique separating equilibrium boils down to proving that $\Theta^* = [\hat{\theta}, \infty)$ for some finite $\hat{\theta}$.

To begin with, this is clearly true if $\ell = 1$: let's say by contradiction that $\exists\theta' < \theta''$ s.t. θ' finds it optimal to disclose, whereas θ'' doesn't. This implies that $b(\theta') > s(\theta') + c(\theta')$ but $b(\theta'') < s(\theta'') + c(\theta'')$ which is ruled out by assumption. The same reasoning also proves equilibrium uniqueness.

More in general, a focal type θ's choice whether to discard $\mathcal{T}$ follows:

$$A^*(\theta) = \text{argmax}_{a\in\{0,1\}} H(a,\theta) = \text{argmax}_{a\in\{0,1\}} a[V(\theta) - c(\theta)] + (1-a)max\{p_{\tau=\emptyset}, s(\theta)\}$$

where

$$\begin{cases} V(\theta) = \int_{x\in X} max\{p_{\tau=x}, r(\theta)\} dG(x|\theta) \\ \\ p_{\tau=x} = E_{\tilde{f}(\theta|x)}[b(\theta)|\theta \le s^{-1}(p_{\tau=x})] \qquad \forall x \in X \\ \\ \tilde{f}(\theta|x) = \frac{g(x|\theta)f(\theta)}{\int_{\Theta^*} g(x|\theta)f(\theta)d\theta} \\ \\ \Theta^* = \{\theta \quad s.t. \quad A^*(\theta) = 1\} \end{cases}$$

If $H(a,\theta)$ satisfies increasing differences, then $A^*(\theta)$ is weakly increasing in θ and there will be a unique separating equilibrium.

$H(a,\theta)$ clearly satisfies increasing differences for $\ell = 1$ and does not satisfy increasing differences for $\ell = 0$. For the proof I am only required to show that if the signal is enough precise, then $H(a,\theta)$ will satisfy increasing differences and the unique separating PBE obtains. This is immediate: assume that $\exists \ell \in (0,1]$ s.t. $H(a,\theta;\ell)$ satisfies increasing differences (and I show that $\ell = 1$, for example, satisfies the requirement). By MIO, it is immediate that $\frac{\partial V(\theta;\ell)}{\partial\theta\partial\ell} \ge 0$, which implies that $H(a,\theta;\ell')$ must satisfy increasing differences as well for any $\ell' \ge \ell$.

Appendix B: Structural estimation of gains from trade

To begin with, I assume that revenue verification is free and perfectly reveals the idea. Free revenue verification implies that the sellers' choice to disclose $\mathbb{1}^{\tau}(\theta, \epsilon_s)$ is solely a function of θ and not of the taste shock. Buyers' bids on unverified ideas cannot reflect unobserved quality by definition: $b_{\mathrm{U}} = p^* + \epsilon_b$. This means that variation in their bids only comes from variation in ϵ_b. This is by assumption orthogonal to θ, so its distribution is identified up to a constant even from selected data of unverified ideas. This allows me do pin down F_ϵ (up to a constant):

$$F_\epsilon(x) = P(\epsilon_b + p^* \leq x + p^*) = P(\epsilon_b + p^* \leq x + p^* | \theta \leq \theta^*) = F(\text{bid}|\text{Unverified Ideas})$$

The fact that they buyers' bid for unverified ideas is unrelated to quality implies that, among verified ideas, the sellers' valuation $s(\theta) + \epsilon_s$ is orthogonal to the buyers' bid. This allows me to identify the distribution of $s(\theta) + \epsilon_s$ by looking at the sale probability, and then back out the distribution of $s(\theta)$ through deconvolution given knowledge of F_ϵ. In other words:

$$\begin{cases} F_{(\epsilon_s + s(\theta)|\theta \leq \theta^*)}(x) = P(\text{successful sale at price x}|\text{Unverified Idea}) \\ \\ \qquad F_{(s(\theta)|\theta \leq \theta^*)}(x) = F_{(\epsilon_s + s(\theta)|\theta \leq \theta^*)}(x) \circledast^{-1} F_{\epsilon_s} \end{cases}$$

Buyers' bids on verified ideas reflect variation in both idea quality and taste shocks: $b_{\mathrm{V}} = b(\theta) + \epsilon_b$. Knowledge of F_ϵ allows me to deconvolute the pdf $f_{b(\theta)}(x|\theta > \theta^*$ from $f_{b_{\mathrm{V}}}$ and f_ϵ. This, however, is the distribution of the buyers' valuation of the "good" ideas only.

The parametric assumption on the proportional gains from trade ($b(\theta)/s(\theta) = k$) allows me to extrapolate the buyers' valuations from the right tail to the left tail of the idea distribution, where no inventor discloses. I estimate k from $\hat{F}_\epsilon$, $\hat{F}_{b(\theta)}$, and the proportion of ideas sold in the verified population (close to 60%). In practice, $P(\text{Sale}|\text{Verified}) = P(\epsilon_b + (1-k)b(\theta) - \epsilon_s \geq 0)$. I draw ϵ_b, ϵ_s and $b(\theta)$ from their empirical distributions and calibrate k to match $P(\text{Sale}|\text{Verified})$ through a grid search.

Welfare when verification is free and available for all[45] is $A = \mathbb{E}[max(\epsilon_s + kb, \epsilon_b + b)]$; welfare when verification is unavailable is $B = \mathbb{E}[1_{\epsilon_s + kb > \epsilon_b + p^*}(\epsilon_s + kb) + (1 - 1_{\epsilon_s + k > \epsilon_b + p^*})(\epsilon_b + b)]$, where the equilibrium price is obtained via Eq. 4. Welfare when the market is shut down completely is $C = \mathbb{E}[\epsilon_s + kb(\theta)]$. Lost gains from trade when verification is unavailable amount to $\frac{A-B}{A-C}$: knowledge of $\hat{F}_\epsilon$, $\hat{F}_b$, $\hat{k}$ allows me to estimate C through Monte Carlo methods.

To conclude, assuming that revenue verification perfectly reveals the truth provides a lower bound on lost gains from trade: if revenue verification is an imperfect signal, the residual asymmetric information between buyers and sellers would drive a wedge between observed bids and true buyers' valuations, and increase welfare losses by Prop. 3.

[45] In this counterfactual, I allow *all* inventors to *freely* verify their revenues. By Proposition 5, in equilibrium all inventors take advantage of this opportunity.