Essays on Economic Theory

by

Tianjiao Dai

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

Doctor of Philosophy in Economics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2019

Author ..... Signature redacted .....
Department of Economics
May 17, 2019

Certified by ..... Signature redacted .....
Glenn Ellison
Gregory K. Palm Professor of Economics
Thesis Supervisor

Certified by ..... Signature redacted .....
Juuso Toikka
Associate Professor of Business Economics and Public Policy
Wharton School of Business, University of Pennsylvania
Thesis Supervisor

Accepted by ..... Signature redacted .....
Ricardo Caballero
Ford International Professor of Economics
Chairman, Departmental Committee on Graduate Studies

Essays on Economic Theory

by

Tianjiao Dai

Submitted to the Department of Economics
on May 17, 2019, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Economics

Abstract

The first chapter considers team incentive schemes that are robust to nonquantifiable uncertainty about the game played by the agents. A principal designs a contract for a team of agents, each taking an unobservable action that jointly determine a stochastic contractible outcome. The game is common knowledge among the agents, but the principal only knows some of the available action profiles. Realizing that the game may be bigger than he thinks, the principal evaluates contracts based on their guaranteed performance across all games consistent with his knowledge. All parties are risk neutral and the agents are protected by limited liability. A contract is said to align the agents' interests if each agent's compensation covaries positively and linearly with the other agents' compensation. It is shown that contracts that fail to do so are dominated by those that do, both in terms of the surplus guarantee under budget balance, and in terms of the principal's profit guarantee when he is the residual claimant. It thus suffices to base compensation on a one-dimensional aggregate even if richer outcome measures are available. The best guarantee for either objective is achieved by a contract linear in the monetary value of the outcome. This provides a foundation for practices such as team-based pay and profit-sharing in partnership.

The second chapter models a ride-sharing market in a traffic network with stochastic ride demands. A monopolistic ride-sharing platform in this traffic network faces a dynamic optimization problem to maximize its per period average payoff in the long run, by choosing policies of setting trip prices, matching ride requests and relocating idle drivers to meet future potential demands. Directly solving the dynamic optimization problem for the ride-sharing platform is computationally prohibitively expensive for a traffic network with reasonably large number of locations and vehicles due to its intrinsic complexity. I provide an theoretical upper bound on the performance of dynamic policies by analyzing a related deterministic problem. Based on the optimal solution to the deterministic problem, I propose implementable heuristic policies for the original stochastic problem that yield average payoffs converging to the theoretical upper bound asymptotically. I also discuss the relative value function iteration method to solve the optimization problem for small-scale markets numerically.

The third chapter examines several discrete-time versions of a dynamic moral hazard in teams problem, a continuous-time model of which has been extensively studied in the pre-

vious literature. The way to transform the continuous-time game into a discrete-time one is not unique, and different discrete-time assumptions with the same continuous-time technology limit lead to different discrete-time equilibria. Regardless of the technology assumption, I find that two-period models can give equilibrium results quite different from that in a continuous-time model: while the continuous-time model predicts existence and uniqueness of symmetric equilibrium, its two-period versions can either have multiple symmetric equilibria or none. Also, not all equilibria in the discrete-time models share features similar to the one predicted by the continuous-time model. The subsequent study of multiple-period models with no learning sheds some light on how the equilibria evolve as the discrete-time model better approximates the continuous-time one.

Thesis Supervisor: Glenn Ellison
Title: Gregory K. Palm Professor of Economics

Thesis Supervisor: Juuso Toikka
Title: Associate Professor of Business Economics and Public Policy
Wharton School of Business, University of Pennsylvania

Acknowledgments

I would like to thank my advisor, Glenn Ellison, for his tremendous help, enlightenment, and support. Glenn not only guided me through numerous crossroads of research directions, but also read through every single draft of the thesis, each time giving extremely detailed and helpful suggestions on how to improve and move forward. He is always ready to help whenever I ask, and has been most understanding during my times of difficulty. This thesis could by no means be completed without his continuous support.

I am incredibly lucky to have Juuso Toikka as my advisor and coauthor. I learned so much from Juuso while working as the teaching assistant in his contract theory class, which later led to the early work of the first chapter of this thesis. I am especially thankful that Juuso sees the value of this work as beyond a simple extension of an existing theory, and collaborating with him to develop it further has become one of my greatest experience in research. I also deeply appreciate our weekly chats during the semester when I first taught 14.01 as an instructor, where I got so many useful tips on teaching as well as on research and life.

I am also deeply indebted to Robert Gibbons, who led me through my transition from coursework to research in my third year. It is with his guidance and encouragement that I finished my first research project, which later turned into the third chapter of this thesis. His lectures in organizational economics has been especially intriguing to me, and has laid the foundation for my later research.

I would like to extend my deepest gratitude to my undergraduate advisor, Yingyi Qian, who led me into the field of economics. Chatting with him has always been a great pleasure and left me with a refreshed mind and new perspectives. His continued mentorship throughout my time at MIT has been invaluable to me.

In addition to my advisors, I thank all faculty members, in or outside of MIT, who has shared with me knowledge and ideas through various courses and seminars. In particular, I am very grateful to Muhamet Yildiz and Bengt Holmstrom, from whom I learned a great deal about economic theory. I am also greatly indebted to Botond Koszegi and Ulrike Malmendier, who taught me so much about behavioral economics during my time in UC Berkeley as an

exchange college student. I would also like to thank all the staff at the department for kindly addressing my questions and requests. Many thanks to Gary King for his continuous and superior support over all these years.

I am fortunate to have many wonderful classmates and friends who made my graduate years an enjoyable and memorable journey of my life. Special thanks go to Mengxi Wu, for all the time we spent together, and Fei Song, for your company in the later years. My thanks also go to Mark Hou, German Gieczewski, Jie Bai, Hongkai Zhang, Harry Di Pei, Sophie Sun, Yufei Wu and many others for their precious friendships and help all along.

Finally, I would like to thank my parents and grandparents for their always unconditional love. It is only by their unwavering support over the last few years that I have been granted the luxury of being able to complete this thesis. I would like to dedicate this dissertation to my husband, Lili Gao. Thank you for your genuine love and tremendous support over all those years, and for always being there when I needed you the most.

Contents

List of Figures

List of Tables

Chapter 1

Robust Incentives for Teams

[1]

1.1 Introduction

Much of economic activity is performed by teams, broadly defined to encompass groups of agents such as partnerships, committees, research groups or start-ups, and work teams in manufacturing and services. The classical contract-theoretic approach to incentivizing such teams, pioneered by Holmström (1982), emphasizes the informational aspects of the problem. It holds that any signal informative of an agent's action should optimally be used to determine his compensation. This leads to contracts that are sophisticated and highly context-dependent. Moreover, there is no reason for compensation to be team-based, unless it is exogenously assumed that the outcome only provides information on the team's aggregate performance. Both predictions are at odds with incentive schemes typically observed in practice, which tend to be simpler and often include team-based pay even if information about individual performance is available. For instance, partnerships commonly operate under a simple profit-sharing agreement.

In this paper, we investigate foundations for such simple incentive schemes by considering contracts that are robust to nonquantifiable uncertainty about the game played by the agents.

[1]This paper is joint work with Juuso Toikka, Wharton School of Business, University of Pennsylvania, 3620 Locust Walk Philadelphia PA 19104, toikka@wharton.upenn.edu.

Our model is based on Holmström's (1982) team production problem, where each agent takes an unobservable action at a private cost, and the profile of actions stochastically determines a contractible outcome that may convey information about both aggregate as well as individual performance. We assume that all parties are risk-neutral and the agents are protected by limited liability, but impose no particular structure on the production technology.

The game is common knowledge among the agents, perhaps by virtue of their expertise, or because it is simply evident now that they have been called to act. However, inspired by Carroll's (2015) work on the foundations of linear contracts in principal-agent problems, we assume that the principal designing the contract only knows some of the actions available to each agent, and hence he only knows some of the action profiles in the game. Realizing that the game may be bigger than he thinks, but not having a prior on the set of possible games, the principal evaluates contracts based on their guaranteed performance across all games consistent with his knowledge.

Our first result shows that guaranteeing good performance either in terms of the expected surplus for a budget-balanced team, or in terms of the principal's own profit if he is the residual claimant, requires that a contract align the agents' interests. In particular, each agent's compensation should covary positively and linearly with the compensation of all other agents. Such a contract has a natural representation in terms of a one-dimensional aggregate of the outcome, the value of which determines everyone's compensation, so we can reasonably interpret the contract as providing team-based compensation. Contracts of this form dominate all other contracts. Thus, team-based compensation is optimal even though rich measures of individual performance may be available.

The necessity of interest alignment derives from the fact that when a contract induces disagreement about the ranking of outcomes among the agents, then—should the game provide the opportunity for it—each agent will seek personal gain at the others' expense. We can then find games where this creates a "race to the bottom," with the unique equilibrium leading to the worst possible outcome. We illustrate the basic intuition in the context of a rank-order tournament after having introduced the model. While the result is reminiscent of Carroll's (2015) linearity result for principal-agent problems, the two are logically distinct: the definition of interest alignment only involves the payments to the agents, so every contract

trivially aligns the agent's interests in the single-agent case.

If a contract that aligns the agents' interests is budget balanced among the agents, then it is in fact a linear contract where each agent is paid a fixed share of the monetary value generated by the contractible outcome. We show that some such linear contract achieves the best possible surplus guarantee within the class of budget-balanced contracts. This provides a possible foundation for profit-sharing agreements in partnerships.

We also show that a linear contract achieves the best possible guarantee for the principal's profit in the case where the principal is the residual claimant for the team's profits and losses. By our first result, the search for the principal-optimal contract can be restricted to contracts that align the agents' interests. Moreover, the candidate optimal contracts can be represented as consisting of a function specifying the agents' total compensation for each outcome, and of shares that determine how it is divided amont the agents. By keeping the shares fixed and focusing on the total compensation, we can adapt Carroll's (2015) argument for the one-agent case to show that the total compensation should be a linear function of the monetary value generated by the outcome (and thus the contract should be linear overall). Heuristically, a contract that aligns the agents' interests ensures that no agent can seek personal gain at the expense of the other agents. Requiring that this not happen at the expense of the principal, either, implies that the agents' compensation must covary linearly with the principal's payoff as well, leading to a linear contract.

Whether the optimal guarantees for surplus and profit are positive depends on the severity of the free-rider problem. Unlike in the case of one agent, it is not enough that some known action profile generate a positive surplus. The condition that characterizes known production technologies for which the optimal guarantees are non-trivial comprises of a virtual surplus calculation: a social planner should be able to generate positive surplus in a model where the agents' costs are appropriately inflated to account for the robustness concern. Thus the theory here predicts that, even absent setup costs, only sufficiently profitable teams are worth forming.

The question of foundations for linear contracts has received a great deal of attention in the one-agent case, starting with Holmström and Milgrom (1987). See Carroll (2015) for a review of this literature. As we focus on the contracts' guaranteed performance, our work

belongs to the literature studying worst-case optimal contracts in various settings—see, for example, Hurwicz and Shapiro (1978), Chung and Ely (2007), Chassang (2013), Frankel (2014), Garrett (2014), Yamashita (2015), Carroll (2017), Carroll and Segal (2017), and Marku and Ocampo Diaz (2017). Similar robustness concerns motivate the work on robust mechanism design following Bergemann and Morris (2005), and the analysis of approximately optimal contracts in locally misspecified models by Madarász and Prat (2017).

Other theoretical explanations have been put forth for the use of profit-sharing, and for the prevalence of partnerships as an organizational form in the professional services industry—see, for example, Garicano and Santos (2004) or Levin and Tadelis (2005). These papers either exogenously restrict the contract space, or solve for the optimal contract in particular parametric models. Che and Yoo (2001) show that team-based compensation can be a part of the optimal mix of formal and relational incentives in a repeated partnership problem where the agents can observe each others' actions. Our work provides a complementary perspective, showing that team-based compensation arises as a robustly optimal contract in a static setting where the agents cannot monitor each other.

Finally, there is an extensive management literature on teams. The result that contracts should align the agents' interests connects with some of the themes in this literature. For example, Hackman (2002) posits that one of the key enabling conditions for work-team effectiveness is the existence of a compelling direction that should specify ends but not means. Interpreting the "means" as the agents' actions and the "ends" as the contractible outcome, a contract that aligns the agents' interests provides just that.[2]

1.2 Model

We consider the problem of a principal incentivizing a team of agents, indexed $i = 1, \ldots, I$. Each agent takes an unobservable action a_i from a finite set A_i at a private cost $c_i(a_i) \geq 0$. The cost can be interpreted as monetary, or as simply describing the agent's preferences over the available actions. The resulting action profile $a = (a_1, \ldots, a_I) \in A := \times_{i=1}^{I} A_i$

[2]This is true quite literally: the parameter d in our Definition 1 is the direction of the ray in $\mathbb{R}_+^I$ along which all payment profiles lie.

determines stochastically the team's observable output y, an element of a finite set Y of possible outcomes. The distribution of y given a is denoted $F(a) \in \Delta(Y)$. We refer to the tuple (A, c, F), where $c = (c_1, \ldots, c_I) : A_1 \times \cdots \times A_I \to \mathbb{R}^I_+$ is the profile of cost functions and $F : A \to \Delta(Y)$ is the family of output distributions, as the *technology*.

The outcome y provides a measure of the team's performance, possibly along multiple dimensions, and serves as a signal about the agents' actions. Its intrinsic monetary value is denoted $v(y)$. For example, $v(y)$ may be the expected market value of the team's production conditional on the signal y, or it may reflect how the principal aggregates different dimensions of performance. We denote by y_0 the least desirable outcome and normalize its value to zero, i.e., $v(y_0) = \min v(Y) = 0$. (y_0 can be chosen arbitrarily among the minimizers if the worst outcome is not unique.) To avoid trivialities, we assume $\max v(Y) > 0$.

The agents do not have preferences over the outcomes per se, but the principal can guide them by designing an incentive scheme that rewards the agents based on the team's output. We assume that the agents are protected by limited liability, meaning that payments to them have to be non-negative. An incentive scheme, or a *contract*, is thus a function $w : Y \to \mathbb{R}^I_+$ that specifies a payment profile $w(y) = (w_1(y), \ldots, w_I(y))$ for every possible outcome $y \in Y$. The net payoff of agent i is then $w_i(y) - c_i(a_i)$, with the principal receiving $v(y) - \sum_i w_i(y)$. All parties are assumed risk neutral.

Given a contract w, the convex hull of all payment profiles is denoted $W := \mathrm{co}(w(Y))$. We say that the contract w is *budget balanced* if the value of output is shared by the agents, i.e., if $\sum_i w_i(y) = v(y)$ for all y.

The principal designs the contract either to maximize total surplus subject to budget balance, or to maximize his profits. However, he does so without full knowledge of the game played by the agents. Specifically, inspired by Carroll (2015), we assume that the technology (A, c, F) is common knowledge among the agents, but the principal only knows some finite set $A^0 = \times_{i=1}^I A_i^0$ of action profiles with an associated profile of cost functions $c^0 : A^0 \to \mathbb{R}^n_+$ and outcome distributions $F^0 : A^0 \to \Delta(Y)$, collectively referred to as the *known technology*. The principal believes that the true technology may be any (A, c, F) such that $A \supseteq A^0$ and $(c, F)|_{A^0} = (c^0, F^0)$. That is, the true technology contains the action profiles known to the principal, and the true costs and output distributions associated with these profiles conform

with the principal's knowledge. (Note that the set of possible outcomes Y is held fixed; it is known by all parties.) To simplify notation, we suppress the cost functions and outcome distributions when this causes no confusion, writing A^0 and A for the known and the true technology, respectively.

Together a contract w and the (true) technology A induce a normal form game $\Gamma(w, A)$ between the agents. We let $\mathcal{E}(w, A)$ denote its set of mixed strategy Nash equilibria. An equilibrium exists because A was assumed finite. In case there are many, we adopt the usual partial-implementation assumption from contract theory and focus on the equilibrium that is best for the principal's objective.[3] Thus, the expected total surplus induced by the contract w given technology A is

$$S(w, A) := \max_{\sigma \in \mathcal{E}(w,A)} \Big(\mathbb{E}_{F(\sigma)}[v(y)] - \sum_a \sigma(a) \sum_i c_i(a_i) \Big),$$

where $F(\sigma)$ is the outcome distribution induced by F and the equilibrium strategy profile σ. Similarly, the principal's expected profit from the contract w given technology A is

$$V(w, A) := \max_{\sigma \in \mathcal{E}(w,A)} \mathbb{E}_{F(\sigma)}[v(y) - \sum_i w_i(y)].$$

Faced with the uncertainty about the game played by the agents, the principal ranks contracts according to their guaranteed expected performance over all possible (finite) technologies. For total surplus and profits, these guarantees are, respectively,

$$S(w) := \inf_{A \supseteq A^0} S(w, A) \quad \text{and} \quad V(w) := \inf_{A \supseteq A^0} V(w, A).$$

We say that a contract is *team-optimal* if it maximizes $S(w)$ over all budget-balanced contracts. A contract is *principal-optimal* if it maximizes $V(w)$. Note that the guaranteed expected surplus satisfies $S(w) \geq -\sum_i \underline{c}_i^0$, where $\underline{c}_i^0 := \min c_i(A_i^0)$, since each agent can ensure a payoff of $-\underline{c}_i^0$ by playing the least-cost action in A_i^0 given any technology $A \supseteq A^0$.

[3]This minimizes the departure from the standard model and ensures the existence of an optimal contract. Essentially the same results obtain under the alternative assumption that the agents play the worst equilibrium for the principal among equilibria that are not strictly Pareto dominated for the agents, but in this case optimal contracts may only exist in the sense of a limit.

Similarly, the *zero contract* $w \equiv 0$ yields a nonnegative expected profit from any technology, and hence $V(0) \geq 0$.

Some remarks regarding the formulation are in order. The most immediate interpretation is that the principal is designing the contract for a single team, not fully aware of the game the agents are playing. For example, this uncertainty may reflect the agents' superior knowledge of the situation. Or it could be due to the principal having to design the contract before the details of the team's operating environment are known, or even who the team's members will be. Importantly, however, the principal can envision and evaluate all possible outcomes that may arise as a result of the teams activities, i.e., he knows the set Y and the mapping $v : Y \to \mathbb{R}$. The fact that Y is held fixed is not restrictive as our main findings do not require the output distributions in the known technology to have full support. Thus Y can contain outcomes that are impossible under the known technology. (We selectively invoke a full-support assumption to state stronger versions of some of our results.)

An alternative interpretation of the model is that the principal is designing a contract to be used in a number of different situations, perhaps by many different teams, and wants the contract to guarantee good expected performance in all of them. In case of the profit guarantee, a concrete example might be a collection of self-managing teams, such as the cabin crews of a large international airline. Each crew may face a multitude of situations at the airport and on board depending on the model and the condition of the aircraft, the number and types of passengers, the weather and possible delays caused by it or by technical problems, and so on. The realized situation may be apparent to all crew members (captured by the agents knowing A), but it may be too difficult and costly to communicate or verify all of this information about the circumstances to a third party for the contract to depend on it. Moreover, large airlines have numerous cabin crews, whose compositions change frequently, so robust performance in a broad range of circumstances may seem desirable.[4]

We have deliberately assumed that the contract can only condition on the outcome y. We view this assumption as capturing the spirit of the robustness exercise. However, as the agents are assumed to know the true technology, a more general contract could first ask the agents to report the true technology A, and then determine the incentive scheme w

[4]See Hackman (2002) and references therein for a discussion and case studies of cabin crews.

based on the reports. With one agent, Carroll (2015) shows that such screening contracts do not alter the findings. But with many agents, there does exist an equilibrium where the agents truthfully report the technology, and—given our partial implementation assumption—the situation reduces to the standard Bayesian case. A stronger notion of implementation together with limited liability would prevent this trivial solution, but a detailed analysis of this case seems difficult, and we leave it for future work.

1.3 Necessity of Interest Alignment

We start the analysis by showing that contracts that fail to align the agents' interests can be easily improved upon. We also show that, under mild additional assumptions about the known technology, interest alignment becomes necessary for a contract to have a nontrivial surplus or profit guarantee.

In order to motivate our definition of interest alignment and to develop intuition for the result, it is useful to first consider a contract under which the agents are in direct competition—a rank-order tournament. To this end, suppose for a moment that there are just two agents and that an outcome is a pair $y = (y_1, y_2)$ listing their outputs. Let Y be any finite grid on $\mathbb{R}^2_+$ containing the origin and at least one outcome with $y_1 > y_2$ and another with $y_2 > y_1$. The value to the principal is the sum $v(y) = y_1 + y_2$, and thus $y_0 = (0, 0)$. A rank-order tournament is a contract specifying three payment levels: $w_i(y) = b > 0$ if $y_i > y_{-i}$, $w_i(y) = b/2$ if $y_i = y_{-i}$, and $w_i(y) = 0$ if $y_i < y_{-i}$. That is, the agent with the highest output gets a bonus b, which is shared equally in case of a tie.

Because of the form of the contract, agent 1 has an incentive to pursue actions that increase the likelihood that his output is greater than agent 2's output, or $y_1 > y_2$. If the only way to do this is by increasing y_1, then the tournament does incentivize the agent to exert effort towards increasing total output.[5] But as pointed out by Lazear (1989), it is also in agent 1's interest to sabotage agent 2 to lower y_2, for example, by refusing to help. He may also try to claim credit for some of agent 2's output—or even outright steal it—to shift

[5]Holmström (1982) showed that for some specific choices of technology, a tournament is the optimal contract for a principal who knows the game played by the agents and designs the contract to maximize his expected payoff. The (sub-)optimality of the tournament in this sense plays no role in the example.

	A_2^0	a_2'
A_1^0	$\cdots$	$-c_1(a_1), b$
a_1'	$b, -c_2(a_2)$	$b/2, b/2$

Figure 1-1. The game $\Gamma(w, A)$ for the tournament example
Since b is the highest feasible payoff, a_i' is a weakly dominant strategy. To see that a' is the unique equilibrium, fix a mixed strategy equilibrium σ. If the support of σ is contained in A^0, then some agent i's expected payoff is at most $b/2$, whereas deviating to a_i' yields b for sure. Hence, a_i' must be in the support of σ_i for some agent i. But then a_{-i}' is the unique best response for agent $-i$, and thus σ_{-i} must play it with certainty. This in turn implies that σ_i must play a_i' with certainty. Therefore, σ is simply the pure-strategy profile a'.

some of y_2 to y_1. To the extent that such actions distract from agent 1's productive efforts, they lead to lower total output. Indeed, if both agents can engage in such activities, even the best equilibrium for the principal may yield no output.

More formally, given any known technology A^0, consider a technology A where each agent has an additional zero-cost action a_i' so that $A_i = A_i^0 \cup \{a_i'\}$ for $i = 1, 2$. The action a_i' results deterministically in some outcome y^i that has agent i winning the tournament if agent $-i$ plays any action in A_{-i}^0 (i.e., $y_i^i > y_{-i}^i$). Think of a_i' as an activity that benefits agent i at the other agent's expense as discussed above. However, suppose that if both agents engage in this activity, then nothing is produced: the profile $a' = (a_1', a_2')$ leads to the outcome $y_0 = (0, 0)$ with certainty. It is easy to verify that a_i' is then a weakly dominant strategy for each agent in the game $\Gamma(w, A)$, and a' is the unique equilibrium—see Figure 1-1.

The principal's profit given technology A is $V(w, A) = v(y_0) - b = -b$, and thus the tournament's profit guarantee is negative: $V(w) \leq V(w, A) < 0$. The principal would be better off offering the zero contract. In fact, as the unique equilibrium of $\Gamma(w, A)$ yields y_0 with certainty, the tournament's profit guarantee would be nonpositive even if rewarding the agents was costless to the principal (i.e., if his payoff was just $v(y)$).

To motivate our definition of interest alignment, it is useful to represent the above argument graphically. In Figure 1-2.a, the line segment between $(b, 0)$ and $(0, b)$ is the convex hull of payment profiles $W = \text{co}(\{(0, b), (b/2, b/2), (b, 0)\})$. The new action a_i' allows agent i to force at zero cost his most preferred point in W (i.e., $(0, b)$ or $(b, 0)$) if agent $-i$ plays any action in A_{-i}^0. As the expected payment profile $\mathbb{E}_{F(\sigma)}[w(y)]$ under any mixed strategy profile σ lies somewhere in W, at least one agent thus has a profitable deviation if the other

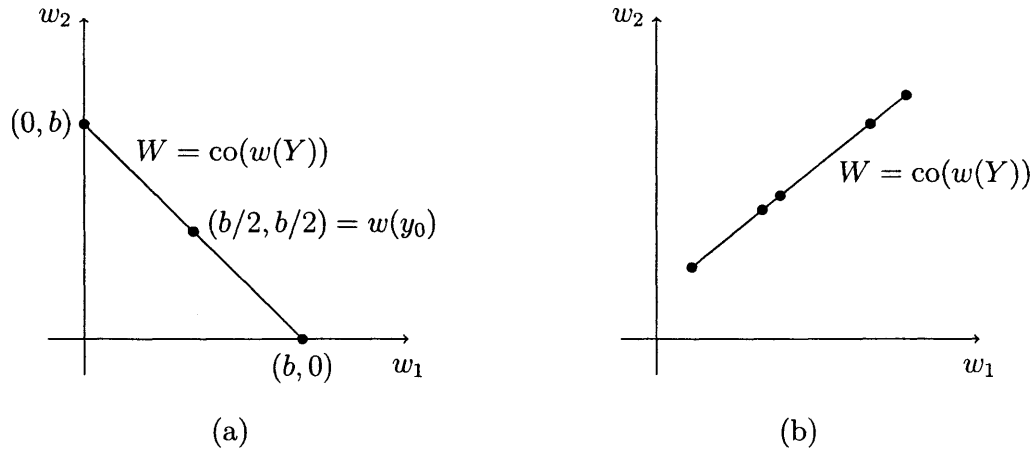


Figure 1-2. (a) The rank-order tournament (b) A contract that aligns the agents' interests

agent's strategy puts positive probability only on known actions. This rules out equilibria with support in A^0. Finally, a' yields rewards $w(y_0) = (b/2, b/2)$ at the midpoint of the line segment; this point is better for each agent than the other agent's most preferred point, so a' is the unique equilibrium.

A rank-order tournament is special in that the agents' interests are in direct conflict: for any outcome distributions $F, G \in \Delta(Y)$, whenever $\mathbb{E}_F[w_1(y)] > \mathbb{E}_G[w_1(y)]$ so that agent 1 prefers F to G, we have $\mathbb{E}_G[w_2(y)] > E_F[w_2(y)]$ so that agent 2 prefers G to F. Graphically, this is equivalent to W being a downward-sloping line segment as in Figure 1-2.a so that the agents have opposite preferences over the points in W. However, it turns out that the same perverse incentives that undermine the tournament can arise in contracts that induce far less disagreement about the desirability of different outcome distributions. To completely rule out such disagreement, W must consist of a (weakly) increasing line segment as in Figure 1-2.b. This suggests the following definition.

Definition 1. A contract w *aligns the agents' interests* if all payment profiles lie on the same ray in $\mathbb{R}^I_+$, i.e., if $w(Y) \subset \{\underline{w} + dt : t \in \mathbb{R}_+\}$ for some $\underline{w}, d \in \mathbb{R}^I_+$.

A contract that does not satisfy the definition is said to *fail to align the agents' interests.* The tournament in Figure 1-2.a is an example. Note that any contract for which the interior of W is non-empty as in Figures 1-3 and 1-4 below fails to align the agents' interests.

Definition 1 is equivalent to the requirement that for all agents i and j and all outcome

distributions $F, G \in \Delta(Y)$, $\mathbb{E}_F[w_i(y)] > \mathbb{E}_G[w_i(y)]$ implies $\mathbb{E}_F[w_j(y)] \geq \mathbb{E}_G[w_j(y)]$. That is, no two agents disagree on the ranking of any pair of outcome distributions, albeit one of the agent's preference may be strict and the other's weak (if the latter is globally indifferent). The equivalence follows by noting that each point in the convex hull of payment profiles W is the expected profile $\mathbb{E}_F[w(y)]$ for some $F \in \Delta(Y)$. Hence, the agents do not disagree on the ranking of distributions precisely when W, and thus $w(Y)$, lies along a ray in $\mathbb{R}^I_+$.

A second equivalent condition is the existence of outcomes $\overline{y}$ and $\underline{y}$ with $w(\overline{y}) \geq w(\underline{y})$ such that, for all $y \in Y$, we have $w(y) = (1-\lambda)w(\underline{y}) + \lambda w(\overline{y})$ for some $\lambda \in [0,1]$. The parameter λ has a natural interpretation as a measure of the team's performance on a scale from zero to one. In this sense a contract that aligns the agents' interests prescribes team-based compensation.

Note that any constant contract satisfies Definition 1. For example, the zero contract aligns the agents' interests. It is also worth noting that the definition only concerns the agents, and so in general it is silent on how the payments relate to the value of the outcome. However, if the contract is also budget balanced, then interest alignment is equivalent to the agents dividing the value $v(y)$ amongst themselves according to some fixed shares.

Lemma 1. *A contract w is budget balanced and aligns the agents' interests if and only if there exists $\alpha = (\alpha_1, \ldots, \alpha_n) \in [0,1]^I$ such that $\sum_i \alpha_i = 1$ and $w_i(y) = \alpha_i v(y)$ for all i and y.*

Proof. Clearly a contract of this form is budget balanced and aligns the agents' interests. For the converse, note that by budget balance we can take $\underline{y} = y_0$ and $\overline{y} \in \arg\max_y v(y)$ in the second equivalent condition above. Fixing y, we sum over i and use budget balance again to get $v(y) = \sum_i w_i(y) = (1-\lambda)\sum_i w_i(y_0) + \lambda \sum_i w_i(\overline{y}) = (1-\lambda)v(y_0) + \lambda v(\overline{y}) = \lambda v(\overline{y})$. Hence, $\lambda = v(y)/v(\overline{y})$. Noting that $w_i(y_0) = 0$ by limited liability and budget balance, we thus have $w_i(y) = (w_i(\overline{y})/v(\overline{y}))v(y)$, so taking $\alpha_i = w_i(\overline{y})/v(\overline{y})$ yields the result. □

Our first main result shows that any contract that fails to align the agents' interests can be easily improved upon regardless of whether we are interested in profits or total surplus.

Theorem 1. *If a contract w fails to align the agents' interests, then $V(w) \leq V(0)$. If, in addition, w is budget balanced, then $S(w) \leq S(w')$ for every contract w' that is budget balanced and aligns the agents' interests.*

That is, the guaranteed expected profit of a contract that fails to align the agents' interests is no better than that of the zero contract, generalizing the observation from the tournament example. And if the contract is also budget balanced, then its guaranteed expected surplus is weakly worse than the guarantee obtained by arbitrarily distributing shares across the agents. These results imply, inter alia, that we can restrict attention to contracts that align the agents' interests when searching for optimal ones.

We prove Theorem 1 by finding for any contract that fails to align the agents' interests a (sequence of) game(s) with poor performance. The construction is more involved, but the basic idea is the same as in the tournament example: misalignment erodes a contract's guaranteed performance because, if given the opportunity, agents will seek personal gain at the expense of others, and this can lead to all equilibria being bad for the principal.

Before turning to the proof, we note that Theorem 1 can be strengthened under additional assumptions about the known technology to show that interest alignment is necessary in order to obtain any nontrivial performance guarantees.

We need the following definitions. An action profile $a \in A^0$ satisfies *full support* if $F(a) \neq \delta_{y_0}$ (where δ_{y_0} is the Dirac measure at y_0) implies that $F(a)$ has full support on Y. It satisfies *costly production* if $\mathbb{E}_{F(a)}[v(y)] > 0$ implies $c_i(a_i) > 0$ for some agent i. The following corollary shows that if each known action profile satisfies either of these, the worst case for any contract that fails to align the agents' interests is that no value is created.

Corollary 1. *Suppose every action profile in the known technology A^0 satisfies full support or costly production. If a contract w fails to align the agents' interests, then there exists a sequence of technologies $A^n \supseteq A^0$ such that*

$$\sup_{\sigma \in \mathcal{E}(w, A^n)} F(\{y \in Y : v(y) > 0\} | \sigma) \to 0.$$

The value of the team's equilibrium output converges to zero as $n \to \infty$, so the principal's profit is nonpositive in the limit (and it would be so even if the principal didn't have to pay the agents' compensation out of pocket). Moreover, the construction in the proof uses actions whose costs are no lower than the costs of the known actions, implying that the equilibrium total surplus converges to its theoretical lower bound. This establishes the following corollary.

Corollary 2. *Under the assumptions of Corollary 1, if a contract w fails to align the agents' interests, then $V(w) \leq 0$ and $S(w) \leq -\sum_i \underline{c}_i^0$.*

Corollary 2 shows that under relatively weak additional assumptions, contracts that fail to align the agents' interests are not only dominated in the sense of Theorem 1; they fail to improve on the trivial bounds both for profits and total surplus. In fact, if the contract is also budget balanced, an even weaker assumption will do.

Corollary 3. *Suppose the known technology A^0 does not contain an action profile a such that $c(a) = 0$ and $\operatorname{supp} F(a) \subseteq \arg\max_y v(y)$. If a budget balanced contract w fails to align the agents' interests, then $S(w) \leq -\sum_i \underline{c}_i^0$.*

That is, unless it is costless to produce the most valuable outcome(s) with certainty under the known technology, any budget balanced contract that fails to align the agents' interests has only the trivial surplus guarantee.

Of course, the above results are silent on whether contracts that do align the agents' interests can improve on the trivial guarantees. We address this question in Sections 1.4 and 1.5, which consider, respectively, team-optimal and principal-optimal contracts.

1.3.1 Proof of Theorem 1

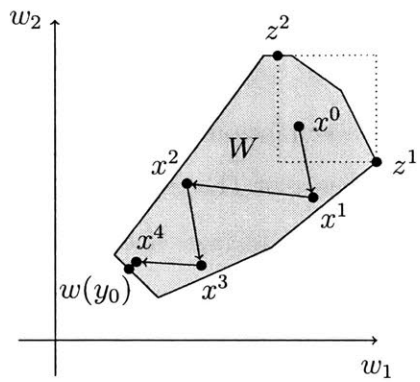
We present here the proof of Theorem 1, relegating those of the corollaries to the Appendix. Throughout this section, fix a contract w that fails to align the agents' interests. Let

$$Y^* := \bigcap_{i=1}^{I} \arg\max_{y \in Y} w_i(y).$$

By definition, any $y \in Y^*$ simultaneously maximizes the payment to every agent. Graphically, this means that $w(y) \geq x$ for every $x \in W$. Note that Y^* may well be empty.

There are three cases to consider, corresponding to the following three lemmas. The first case is that the set Y^* is empty. This case is similar to the tournament example in that then no single point in W is the best point for all agents.

Lemma 2. *If $Y^* = \emptyset$, then there exists a sequence of technologies $A^n \supseteq A^0$, with unique equilibrium distributions $F^n \in \Delta(Y)$ and $\min c_i(A_i^n) = \underline{c}_i^0$ for all i, such that $F^n \to \delta_{y_0}$.*



	A_2^0	a_2^1	a_2^2	a_2^3
A_1^0	$\cdots$	$\mathbf{z^2}$	u^0	u^0
a_1^1	$\mathbf{z^1}$	$\mathbf{x^0}$	u^1	u^1
a_1^2	u^0	$\mathbf{x^1}$	$\mathbf{x^2}$	u^2
a_1^3	u^0	u^1	$\mathbf{x^3}$	$\mathbf{x^4}$

Figure 1-3. A contract that fails to align the agents' interests with $Y^* = \emptyset$
The gray area is W, the convex hull of payment profiles. The payoff matrix represents the game $\Gamma(w, A)$ in the proof sketch for Lemma 2.

The proof of this and that of the next lemma make use of the fact that the agents' payoffs depend on the outcome distribution $F(a) \in \Delta(Y)$ only through the expected payment profile $\mathbb{E}_{F(a)}[w(y)] \in W$. Conversely, any $x \in W$ is the expected payment profile for some $F \in \Delta(Y)$. Therefore, when constructing a technology A, as far as the agents' incentives are concerned, it suffices to specify the expected payment profiles $x(a) \in W$, $a \in A$.

Proof Sketch. For simplicity, we sketch the proof for the case of two agents, assuming further that the lowest known cost is zero for each agent (i.e., $\underline{c}_i^0 = 0$, $i = 1, 2$).

When $Y^* = \emptyset$ and $I = 2$, the set W is either a downward-sloping line segment as in Figure 1-2.a, or it has a non-empty interior as in Figure 1-3. The first case essentially reduces to the tournament example, so we focus on Figure 1-3.

Consider a technology A where $A_i = A_i^0 \cup \{a_i^1, a_i^2, a_i^3\}$ with $c_i(a_i^k) = 0$ for all i and k (so that each new action is a least-cost action) and where any $a \in A$ involving new actions is assigned an expected payment profile as specified in the right panel of Figure 1-3.

Note that z^i is (one of) agent i's most preferred point(s) in W. Such points z^1 and z^2 are necessarily distinct when $\arg\max_y w_1(y) \cap \arg\max_y w_2(y) = \emptyset$. Taking z^i to be the expected payment profile if agent i plays a_i^1 and the other agent plays any $a_{-i} \in A_{-i}^0$ eliminates equilibria in known actions (i.e., with support in A^0): in any such equilibrium some agent i would necessarily get less than z^i, and hence he could profitably deviate to a_i^1.

If $w(y_0)$ was in the dotted rectangle in Figure 1-3, we could then set $x(a_1^1, a_2^1) = w(y_0)$

with no need for actions a_i^2 and a_i^3. Then (a_1^1, a_2^1) would be the unique equilibrium of the game $\Gamma(w, (A_1^0 \cup \{a_1^1\}) \times (A_2^0 \cup \{a_2^1\}))$, similarly to the tournament example. However, when $w(y_0)$ lies outside the rectangle, as depicted here, this no longer works as at least one agent i prefers the other agent's most preferred point z^{-i} to $w(y_0)$. (They both do in Figure 1-3.)

Instead, we choose a sequence $(x^0, \ldots, x^4)$ in W as in Figure 1-3. That is, each agent i prefers x^0 to z^{-i}, and given any two adjacent elements of the sequence $(x^0, \ldots, x^4)$, agent 1 strictly prefers the odd one and agent 2 the even one. The last element, x^4, is chosen in the interior of W so that we can populate the remaining cells in the payoff matrix in Figure 1-3 with points $u^0 < u^1 < u^2$ such that $u^l < x^k$ for all l, k. (u^i are not shown in the left panel; they can be chosen in the gap between x^4 and $w(y_0)$ if x_2^3 is close enough to x_2^4.)

When the row(s) and column(s) corresponding to A_1^0 and A_2^0 have been eliminated—they are not necessarily strictly dominated, but no agent will play them with positive probability in any equilibrium—the remaining matrix is by construction dominance solvable, with (a_1^3, a_2^3) the unique outcome. Letting $x^4 \to w(y_0)$ thus yields a sequence of technologies whose unique equilibrium expected payment profiles converge to $w(y_0)$, and thus the distributions generating them can be taken to converge to δ_{y_0} as desired.

Note that the number of steps in the path $(x^0, \ldots, x^4)$, and hence the number of actions in the technology A, is dictated by the shape of W, with a narrower set requiring more steps. However, any points x and $x' < x$ in the interior of W can be connected by such a path. □

If $Y^* \neq \emptyset$, the projection of W to some pair of agents' payments is of the form depicted in Figure 1-4. (The interior (relative to $\mathbb{R}_+^2$) is non-empty, since otherwise w would align the agents' interests.) There no longer exist distinct most preferred points for the agents that got the construction in the proof of Lemma 2 started. However, if no known zero-cost action profile maps deterministically to the point z in Figure 1-4, then we can still eliminate equilibria involving profiles in A^0 and drive the outcome to y_0 with essentially the same construction.

Lemma 3. *Let $Y^* \neq \emptyset$. Suppose that, for all $a \in A^0$, $\operatorname{supp} F(a) \subseteq Y^*$ implies $c(a) \neq 0$. Then there exists a sequence of technologies $A^n \supseteq A^0$, with unique equilibrium distributions $F^n \in \Delta(Y)$, such that $\min c_i(A_i^n) \to \underline{c}_i^0$ for all i and $F^n \to \delta_{y_0}$.*

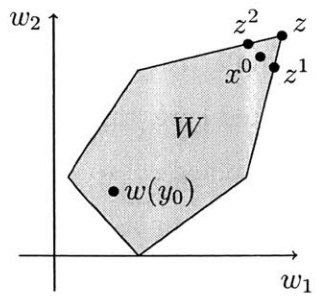


Figure 1-4. A contract that fails to align the agents' interests with $Y^* \neq \emptyset$

Proof Sketch. We again assume for simplicity that $I = 2$ and $\underline{c}_i^0 = 0$, $i = 1, 2$.

Consider a technology A that assigns one new zero-cost action a_i^1 to each agent so that $A_i = A_i^0 \cup \{a_i^1\}$. Choose x^0, z^1, z^2 as in Figure 1-4, i.e., x^0 is in the interior of W and $z_i^i > x_i^0 > z_i^{-i}$. Let the expected payment profile be z^i if only agent i plays the new action a_i^1; let it be x^0 if both agents play the new action.

The profile a^1 is an equilibrium of the game $\Gamma(w, A)$, because a_i^1 is the unique best-response to a_{-i}^1 by construction. In fact, for x^0 close enough to z, it is the only equilibrium. To see this, choose x^0 close enough to z such that

$$x_1^0 + x_2^0 > \mathbb{E}_{F(a)}[w_1(y)] - c_1(a_1) + \mathbb{E}_{F(a)}[w_2(y)] - c_2(a_2) \quad \forall a \in A^0.$$

This is possible, because by assumption every $a \in A^0$ with $\mathbb{E}_{F(a)}[w(y)] = z$ has some agent playing a costly action, and A^0 is finite. The inequality implies that for all $a \in A^0$, we have $z_i^i > x_i^0 > \mathbb{E}_{F(a)}[w_i(y)] - c_i(a_i)$ for some agent i, who thus can profitably deviate to a_i^1. This rules out other pure-strategy equilibria. With some more work one can establish a^1 as the unique mixed-strategy equilibrium as well.

Having escaped the point z, we can now add actions $\{a_i^2, \ldots, a_i^K\}$ to the technology A and use the construction in the proof of Lemma 2 to drive the equilibrium outcome to y_0. □

Finally, there remains the possibility that some known action profile $a^* \in A^0$ ensures that the outcome is in Y^* at no cost to the agents. Then a^* is an equilibrium for any technology $A \supseteq A^0$, and hence the contract w can potentially give a nontrivial profit or surplus guarantee. But a^* is also an equilibrium under the zero contract as well as under

any budget balanced contract that aligns the agents' interests; such contracts can be shown to improve upon w. More precisely, we have the following lemma.

Lemma 4. *Suppose there exists $a \in A^0$ such that $\operatorname{supp} F(a) \subseteq Y^*$ and $c(a) = 0$. Then $V(w) < V(0)$. Moreover, if w is budget balanced, then $S(w) \leq S(w')$ for every contract w' that is budget balanced and aligns the agents' interests.*

We note for future reference that this lemma holds also for any contract w that aligns the agents' interests, different from the zero contract.

Proof. Let $a^* \in A^0$ satisfy the assumption in the lemma. Consider a technology A where $A_i = A_i^0 \cup \{a_i'\}$ and $c_i(a_i') = 0$ for all i. Let $F(a_i', a_{-i}) = F(a^*)$ for all $a_{-i} \in A_{-i}$. Then each agent can ensure his highest feasible payoff $\max w_i(Y)$ by playing a_i'. This implies that any equilibrium $\sigma \in \mathcal{E}(w, A)$ can assign positive probability only to a such that $c(a) = 0$ and $\operatorname{supp} F(a) \subseteq Y^*$. Hence,

$$\begin{aligned} V(w, A) &\leq \max_{a \in A \,:\, c(a)=0 \text{ and } \operatorname{supp} F(a) \subseteq Y^*} \mathbb{E}_{F(a)}\Big[v(y) - \sum_i w_i(y)\Big] \\ &= \max_{a \in A^0 \,:\, c(a)=0 \text{ and } \operatorname{supp} F(a) \subseteq Y^*} \mathbb{E}_{F(a)}\Big[v(y) - \sum_i w_i(y)\Big] < \max_{a \in A^0 \,:\, c(a)=0} \mathbb{E}_{F(a)}[v(y)] \leq V(0). \end{aligned}$$

Above, the second line follows from the first one, since the set of distributions associated with zero-cost profiles is the same in A and A^0; the strict inequality follows, since $w_i(y) > 0$ for $y \in Y^*$ for some agent i because w is different from the zero contract; the last inequality follows since every $a \in A^0$ with $c(a) = 0$ is an equilibrium under the zero contract given any $A \supseteq A^0$. Thus, $V(w) \leq V(w, A) < V(0)$, establishing the first part of the lemma.

For the second part, suppose that w is budget balanced so that $\sum_i w_i(y) = v(y)$ for all y. Then $Y^* \subseteq \arg\max_{y \in Y} \sum_i w_i(y) = \arg\max_{y \in Y} v(y)$. By assumption, there thus exists $a^* \in A^0$ such that $c(a^*) = 0$ and $\operatorname{supp} F(a^*) \subseteq \arg\max_{y \in Y} v(y)$. We claim that $a^* \in \mathcal{E}(w', A)$ for any budget-balanced contract w' that aligns the agents' interests and any technology $A \supseteq A^0$. Indeed, w_i' is of the form $w_i'(y) = \alpha_i v(y)$ for some $\alpha_i \geq 0$ by Lemma 1. So a^* gives each agent his highest feasible payoff under w' as it maximizes $v(y)$ at zero cost. Hence, a^* is an equilibrium and $S(w') \geq \mathbb{E}_{F(a^*)}[v(y)] = \max v(Y) \geq S(w)$. □

Theorem 1 follows from Lemmas 2–4 by noting that each of Lemmas 2 and 3 implies that the profit guarantee satisfies $V(w) \leq V(w, A^n) \leq \mathbb{E}_{F^n}[v(y)] \to 0$, and the surplus guarantee satisfies $S(w) \leq S(w, A^n) \leq \mathbb{E}_{F^n}[v(y)] - \sum \min c_i(A_i^n) \to -\sum \underline{c}_i^0$.

To prove the stronger results in Corollaries 1–3, it suffices to show that the additional assumptions allow us to strengthen Lemma 4. Heuristically, they ensure that there is no known zero-cost action profile $a \in A^0$ such that $\operatorname{supp} F(a) \subseteq Y^*$, or, if one exists, that the outcomes in Y^* have value zero. See the Appendix for details.

1.4 Team-Optimal Contracts

We now consider team-optimal contracts, i.e., contracts that give the best possible expected surplus guarantee $S(w)$ subject to budget balance. By the preceding analysis, it is without loss of generality to restrict attention to contracts that align the agents' interests. We show here that an optimal contract of this form exists, and give a necessary and sufficient condition for it to yield a non-trivial surplus guarantee. We also derive a necessary and sufficient condition for the guarantee to be positive so that the team is worth forming, and discuss how to find the optimal shares.

The following result collects our main findings on team-optimal contracts. For the statement, say that a contract is *linear* if each agent is paid some fixed share $\alpha_i \in [0, 1]$ of the value $v(y)$ for every $y \in Y$. Under budget balance this is equivalent to the contract aligning the agents' interests by Lemma 1.

Theorem 2. *(i) There exists a linear team-optimal contract.*

(ii) A team-optimal contract w guarantees non-trivial expected surplus (i.e., $S(w) > -\sum_i \underline{c}_i^0$) if and only if the known technology A^0 satisfies

$$\max_{a \in A^0} \Big(\mathbb{E}_{F(a)}[v(y)] - \sum_i c_i(a_i) - 2 \sum_{i,j:i \neq j} \sqrt{(c_i(a_i) - \underline{c}_i^0)(c_j(a_j) - \underline{c}_j^0)} \Big) > -\sum_i \underline{c}_i^0. \quad (1.4.1)$$

(iii) If no known action profile $a \in A^0$ satisfies both $c(a) = 0$ and $\mathbb{E}_{F(a)}[v(y)] = \max v(Y)$, then every team-optimal contract that guarantees non-trivial expected surplus is linear.

Part (i) of Theorem 2 implies that profit-sharing is a robustly optimal arrangement for a partnership absent a principal who could act as a sink or a source of funds. Part (ii) gives a necessary and sufficient condition on the known technology A^0 for the optimal surplus guarantee to be non-trivial. The first two terms on the left-hand side of (1.4.1) correspond to the expected surplus. The presence of the third term means that a non-trivial surplus guarantee requires the known technology to be sufficiently productive; simply being able to generate surplus in excess of the agents' minimum costs is not enough. As we explain below, this can be seen as a manifestation of the familiar free-rider problem. Finally, part (iii) gives a sufficient condition for all team-optimal contracts to be linear when (1.4.1) holds. The condition is weak enough that we expect it to hold in most cases of interest.

Part (iii) of Theorem 2 follows immediately from Corollary 3. As for part (i), since budget-balanced linear contracts outperform other budget-balanced contracts by Theorem 1, it suffices to show the existence of an optimal contract within the class of linear contracts. For this we identify the space of such contracts with the compact set $\{\alpha \in [0,1]^I : \sum_i \alpha_i = 1\}$ by Lemma 1. We then use the upper hemi-continuity of the Nash equilibrium correspondence and our equilibrium selection to show that the surplus guarantee $S(w)$ is an upper semi-continuous function of α. As such it achieves a maximum at some α^*; this α^* is a team-optimal contract. See Lemma 1.8.4 for details.

To prove part (ii) of Theorem 2, and to describe the properties of linear team-optimal contracts further, we characterize the guaranteed expected surplus $S(w)$ for any budget-balanced linear contract. Fix any such contract w. Define $U^0(w) \in \mathbb{R} \cup \{-\infty\}$ by setting

$$U^0(w) := \sup_{a \in A^0} \left(\mathbb{E}_{F(a)}[v(y)] - \sum_i \frac{c_i(a_i)}{\alpha_i} \right), \tag{1.4.2}$$

where $0/0 = 0$ and $x/0 = \infty$ for $x > 0$ by convention. We shall see below that $U^0(w)$ can be interpreted as a type of virtual surplus. It turns out to be a key object in the determination of the guarantee $S(w)$. One can easily verify that $U^0(w) = -\infty$ if and only if $\alpha_i = 0$ for some agent i with positive lowest known cost $\underline{c}_i^0 > 0$. Thus, every agent's share being positive, or every agent's lowest known cost being zero, is sufficient for $U^0(w)$ to be finite. In fact, in the latter case we have $U^0(w) \geq 0$, since the sum in (1.4.2) can then be made zero by choosing

a profile of zero-cost actions.

We note first that if $U^0(w) = -\infty$, then we obtain the trivial guarantee $S(w) = -\sum_i \underline{c}_i^0$. Indeed, then $\alpha_i = 0$ for some agent i with $\underline{c}_i^0 > 0$. Consider a technology where this agent i has a new action that costs slightly less than $\underline{c}_i^0$, but leads to the outcome y_0 with certainty. Agent i must play this new action in all equilibria as he has no stake in the outcome, and the other agents must play their least-cost actions as they cannot affect the outcome. Thus all equilibria yield the lowest feasible surplus with certainty.[6] This shows that we can at least weakly improve upon any contract w with $U^0(w) = -\infty$.

For the case $U^0(w) > -\infty$ we have the following characterization.

Lemma 5. *Let w be a budget-balanced contract that aligns the agents' interests. Suppose $U^0(w) > -\infty$. Then*

$$S(w) = \min_{E \in [0, \max v(Y)],\, b \in \mathbb{R}_+^I} E - \sum_i b_i \quad \textit{subject to}$$

$$E - \sum_i \frac{b_i}{\alpha_i} \geq U^0(w), \tag{1.4.3}$$

$$\alpha_i E - b_i \geq -\underline{c}_i^0 \quad \forall i = 1, \ldots, I. \tag{1.4.4}$$

If (E, b) achieves the minimum, then it satisfies (1.4.3) with equality. Furthermore, there exists a minimizer (E, b) with $E = \max\{U^0(w), 0\}$, $Eb_i = 0$, and $b_i \leq \underline{c}_i^0$ for all i.

The interpretation of the above minimization problem is that we are trying to construct a worst-case technology $A \supseteq A^0$ where the best equilibrium $a \in A$ results in the expected value of output $E = \mathbb{E}_{F(a)}[v(y)]$ and cost $b_i = c_i(a_i)$ for each agent i.

The proof of Lemma 5 uses the fact that w induces a potential game between the agents.[7] That $S(w)$ is not less than the minimum follows by verifying that any technology has an equilibrium satisfying (1.4.3) and (1.4.4). To see this, fix any technology $A \supseteq A^0$. Suppose

[6] More formally, consider $A \supseteq A^0$ where $A_i = A_i^0 \cup \{a_i'\}$ for some agent i with $\alpha_i = 0$ and $\underline{c}_i^0 > 0$, and $A_j = A_j^0$ for $j \neq i$. Let $c_i(a_i') = \eta \underline{c}_i^0$ for some $\eta \in (0, 1)$, and let $F(a) = \delta_{y_0}$ whenever $a_i = a_i'$. Because $\alpha_i = 0$, agent i plays a_i' in all equilibria of $\Gamma(w, A)$. Since the other agents cannot affect the outcome, they must play their least-cost actions in A^0. Thus, $S(w) \leq S(w, A) = -\eta \underline{c}_i^0 - \sum_{j \neq i} \underline{c}_j^0 \to -\sum_j \underline{c}_j^0$ as $\eta \to 1$.

[7] All concepts and results related to potential games used in the analysis can be found in Monderer and Shapley (1996). It can be shown that any contract that aligns the agents' interests induces a potential game between the agents, but we do not need the general form of this result.

for simplicity that each agent's share α_i is positive under w. Define the function $P : A \to \mathbb{R}$ by

$$P(a) := \mathbb{E}_{F(a)}[v(y)] - \sum_i \frac{c_i(a_i)}{\alpha_i}.$$

Denoting agent i's payoff by $u_i(a) := \mathbb{E}_{F(a)}[\alpha_i v(y)] - c(a_i)$ we have, for every a_i, a_i', and a_{-i},

$$u_i(a_i, a_{-i}) - u_i(a_i', a_{-i}) = \alpha_i(P(a_i, a_{-i}) - P(a_i', a_{-i})).$$

That is, the function P is a weighted potential for the game $\Gamma(w, A)$. This implies that any $a^* \in \arg\max_{a \in A} P(a)$ is a pure strategy Nash equilibrium. Any such equilibrium a^* satisfies

$$\mathbb{E}_{F(a^*)}[v(y)] - \sum_i \frac{c_i(a_i^*)}{\alpha_i} = \max_{a \in A} P(a) \geq \max_{a \in A^0} P(a) = U^0(w),$$

where the inequality follows because $A \subseteq A^0$, and the last equality is by definition of $U^0(w)$. Thus, there always exists an equilibrium where the expected value of output and costs satisfy (1.4.3). Moreover, we have $\alpha_i \mathbb{E}_{F(a^*)}[v(y)] - c_i(a_i^*) \geq -\underline{c}_i^0$, as otherwise agent i would deviate to the least-cost action in A_i^0, and thus this equilibrium also satisfies (1.4.4).

We generalize the above argument in the Appendix to the case where $\alpha_i = 0$ for some agent(s). Then P is no longer a potential, but the game can still be shown to be a generalized ordinal potential game. We state the general version here for future reference.

Lemma 6. *Let w be a budget-balanced contract that aligns the agents' interests. For every technology $A \supseteq A^0$, there exists a pure-strategy Nash equilibrium $a^* \in \mathcal{E}(w, A)$ such that*

$$\mathbb{E}_{F(a^*)}[v(y)] - \sum_i \frac{c_i(a_i^*)}{\alpha_i} \geq U^0(w),$$

and $\alpha_i \mathbb{E}_{F(a^*)}[v(y)] - c_i(a_i^*) \geq -\underline{c}_i^0$ for all i. ($-\infty \geq -\infty$ is allowed in the first inequality.)

To prove the other direction—that $S(w)$ is not greater than the minimum in Lemma 5—we construct a sequence of technologies $A^n \supseteq A^0$ such that $S(w, A^n)$ converges to the minimum. The idea is the easiest to illustrate when every agent's share is positive, so suppose this is the case. Let (E, b) achieve the minimum in Lemma 5. Consider a technology A that assigns one new action a_i' to each agent. Let $\mathbb{E}_{F(a')}[v(y)] = E$ and $c_i(a_i') = b_i$.

We note first that it is easy to complete the description of A such that the profile of new actions a' is an equilibrium by considering the potential P. To see this, note that (1.4.3) says that $P(a') \geq U^0(w) = \max_{a \in A^0} P(a)$. Thus, if we set $F(a) = \delta_{y_0}$ for every $a \notin A^0 \cup \{a'\}$, then $P(a) = -\sum c_i(a_i)/\alpha_i \leq -\sum b_i/\alpha_i \leq P(a')$ for all such a, since we can assume that the minimizer satisfies $b_i \leq \underline{c}_i^0$ by Lemma 5. Then a' maximizes the potential on A and hence it is an equilibrium with the desired surplus $E - \sum b_i$.

However, in general the equilibrium that maximizes the potential need not be the one with the highest surplus. To rule out equilibria with a higher surplus, we modify the construction to ensure that a' is the unique equilibrium. Roughly, we slightly increase the payoff to the profile a' and then use the gap between $P(a')$ and $U^0(w)$ so created to carefully construct the distributions $F(a)$ for actions $a \notin A^0 \cup \{a'\}$ to eliminate all other equilibria. We relegate the details to the Appendix. Our discussion of the free-rider problem below contains a heuristic derivation of the worst-case surplus under particular assumptions, which essentially provides a non-technical outline of the proof for that case.

We are now in a position to complete the proof of Theorem 2.

Proof of Theorem 2.(ii). Let w be a linear team-optimal contract. Without loss of generality, we may assume $U^0(w) > -\infty$ so that Lemma 5 applies. Let (E, b) be a minimizer. Suppose that w only has the trivial guarantee $S(w) = -\sum_i \underline{c}_i^0 \leq 0$. Then (1.4.3) implies

$$0 \geq S(w) = E - \sum_i b_i \geq E - \sum_i \frac{b_i}{\alpha_i} \geq U^0(w).$$

By Lemma 5 we can then take $E = 0$ and $b_i \leq \underline{c}_i^0$, and hence we must have $b_i = \underline{c}_i^0$ for all i. Since (1.4.3) binds, this in turn implies that $-\sum_i \underline{c}_i^0/\alpha_i = U^0(w)$, or equivalently, that

$$\max_{a \in A^0} \left(\mathbb{E}_{F(a)}[v(y)] - \sum_i \frac{c_i(a_i) - \underline{c}_i^0}{\alpha_i} \right) = 0. \tag{1.4.5}$$

Conversely, if (1.4.5) holds, then $(E, b) = (0, (\underline{c}_1^0, \ldots, \underline{c}_I^0))$ is feasible in the minimization problem, and thus $S(w) \leq -\sum_i \underline{c}_i^0$. We conclude that (1.4.5) is a necessary and sufficient condition for w to yield the trivial guarantee $S(w) = -\sum_i \underline{c}_i^0$.

The left-hand side of (1.4.5) is non-negative, since the sum can be made equal to zero by

choosing a profile of least-cost actions. Hence, w has a non-trivial surplus guarantee if and only if the maximum is positive. Maximizing with respect to α we obtain a necessary and sufficient condition for this to be true for some (and hence also for the optimal) α:

$$\max_{a \in A^0} \max_{\alpha \in [0,1]^I : \sum_i \alpha_i = 1} \left(\mathbb{E}_{F(a)}[v(y)] - \sum_i \frac{c_i(a_i) - \underline{c}_i^0}{\alpha_i} \right) > 0. \tag{1.4.6}$$

For any $a \in A^0$, the inner maximum in (1.4.6) is achieved by setting

$$\alpha_i(a) = \frac{\sqrt{c_i(a_i) - \underline{c}_i^0}}{\sum_j \sqrt{c_j(a_j) - \underline{c}_j^0}},$$

with $0/0 = 1/I$ by convention. Substituting these shares back into (1.4.6) and rearranging then yields the expression in (1.4.1) and establishes part (ii) of Theorem 2. □

1.4.1 Positive guarantee and optimal shares

Part (ii) of Theorem 2 characterizes all known technologies for which a team-optimal contract gives a non-trivial surplus guarantee. However, if the lowest known cost $\underline{c}_i^0$ is positive for some of the agents, this guarantee may still be negative. One possible interpretation of such a positive lowest cost is that the agent has to incur a fixed cost. With this in mind, it is natural to ask when does a team-optimal contract achieve a positive surplus guarantee $S(w) > 0$ so that the fixed cost is worth incurring and the team is worth forming.

The properties of the minimizers in Lemma 5 immediately imply the following result.

Lemma 7. *Let w be a budget-balanced contract that aligns the agents' interests. We have $S(w) > 0$ if and only if $U^0(w) > 0$. Furthermore, if $U^0(w) \geq 0$, then $S(w) = U^0(w)$.*

Therefore, $S(w) > 0$ for some contract w (and thus also for the optimal contract) if and only if $U^0(w)$ can be made strictly positive by maximizing it with respect to α, or

$$\max_{\alpha \in [0,1]^I : \sum_i \alpha_i = 1} U^0(w) = \max_{a \in A^0} \max_{\alpha \in [0,1]^I : \sum_i \alpha_i = 1} \left(\mathbb{E}_{F(a)}[v(y)] - \sum_i \frac{c_i(a_i)}{\alpha_i} \right) > 0.$$

Analogously to (1.4.6), the inner maximum is achieved for any $a \in A^0$ by setting

$$\alpha_i(a) = \frac{\sqrt{c_i(a_i)}}{\sum_j \sqrt{c_j(a_j)}}, \tag{1.4.7}$$

with $0/0 = 1/I$ by convention. Substituting these shares back into the objective establishes the following result.

Theorem 3. *A team-optimal contract guarantees positive expected surplus if and only if*

$$\max_{a \in A^0} \left(\mathbb{E}_{F(a)}[v(y)] - \sum_i c_i(a_i) - 2 \sum_{i,j:i \neq j} \sqrt{c_i(a_i) c_j(a_j)} \right) > 0. \tag{1.4.8}$$

If (1.4.8) *holds and w is a team-optimal contract, then $S(w)$ equals the left-hand side of* (1.4.8). *Moreover, a linear team-optimal contract can then be found by substituting any maximizer from* (1.4.8) *into* (1.4.7).

Analogously to (1.4.1), the left-hand side of condition (1.4.8) includes an extra term in addition to the expected surplus, implying that the maximized expected surplus from the known technology has to be sufficiently positive for the guarantee to be positive. Of course, (1.4.1) and (1.4.8) coincide if the lowest known cost is zero for each agent.

Theorem 3 shows that when the optimal surplus guarantee is positive, a linear team-optimal contract can be found by first solving the maximization problem in (1.4.8) to obtain a maximizer a^*, and then backing out the shares $\alpha_i^* = \alpha_i(a^*)$ using (1.4.7). This provides a way to solve for an optimal contract in parametric examples. The simplest such example is found by taking the known technology to be a singleton, in which case the optimal shares are just (1.4.7) evaluated at the only known action profile.

Even without any additional structure, it is immediate from (1.4.7) that an agent's share should be larger, the costlier is his action in the profile a^*. However, in general the optimal shares reflect more than just the agents' costs, because a^* depends also on how the actions affect the distribution of outcomes.

For completeness, we note that if (1.4.1) holds, but (1.4.8) is not satisfied, then a linear team-optimal contract can still be found by maximizing $S(w)$ with respect to α using the

characterization in Lemma 5. But this is a tedious exercise at best. Finally, if (1.4.1) is not satisfied, any contract just gives the trivial guarantee $-\sum \underline{c}_i^0$, so any shares are optimal.

1.4.2 On the free-rider problem

In order to see the economic intuition behind the surplus guarantee, it is useful to consider the case where the lowest known cost is zero for each agent (i.e., $\underline{c}_i^0 = 0$, $i = 1, \ldots, I$) so that $U^0(w)$ is nonnegative. The surplus guarantee from any linear budget-balanced contract w is then simply $S(w) = U^0(w)$ by Lemma 7.

The surplus guarantee $U^0(w)$ can be interpreted as a type of *virtual surplus*: it accounts for the incentive and robustness concerns by inflating the agents' costs with their shares. Its roots are in the free-rider problem. Namely, if there is only one agent, then the virtual surplus equals the true surplus as the agent receives the full value of output. But with two or more agents, it is impossible to promise the full value to every agent—captured by some of the shares α_i being less than one—resulting in the virtual surplus being lower than the true surplus. As is evident from the derivation of Theorem 3 above, the third term in (1.4.8) captures what is left of this effect after the shares have been optimized. It is also responsible for the extra term in (1.4.1).

The following thought experiment explains the functional form of $U^0(w)$. Suppose that the known technology consists of just one action profile: $A^0 = \{a^0\}$. Consider trying to lower the expected surplus relative to a^0 by giving agent 1 a new lower-cost action a_1'. To keep agent 1 indifferent and thus willing to play the new action, we can lower the expected value of output by $(c_1(a_1^0) - c_1(a_1'))/\alpha_1$. If $\alpha_1 < 1$, this reduces the total surplus as the reduction in the expected value of output is larger than agent 1's cost saving. That is, the deviation by agent 1 imposes a negative externality on the other agents; this is precisely the free-rider problem. Letting $c_1(a_1') = 0$ we obtain the maximal reduction of $c_1(a_1^0)/\alpha_1$ in the expected value of output. We can then give agent 2 a zero-cost action a_2' to obtain a further reduction of $c_2(a_2^0)/\alpha_2$, and so on. Continuing this process, we obtain a zero-cost action profile a' whose expected value of output is given by $\mathbb{E}_{F(a')}[v(y)] = \mathbb{E}_{F(a^0)}[v(y)] - \sum c_i(a_i^0)/\alpha_i$. As the costs are zero, this is also the expected surplus from the profile a'.[8] By inspection of (1.4.2), we

[8]This is essentially what the worst-case technology used in the proof of Lemma 5 reduces to in this case,

have thus arrived at the formula for $U^0(w)$ for the case of one known action profile.

Heuristically, when there are multiple known action profiles, the maximum in (1.4.2) identifies the one for which the above process yields the highest remaining surplus.

Two further remarks are in order: First, if some agents have fixed costs so that $\underline{c}_i^0 > 0$, the above logic still applies. Indeed, as long as $U^0(w) \geq 0$, the worst-case can still be obtained with zero-cost actions—that is, Lemma 5 has a minimizer with $b = 0$—and hence we have $S(w) = U^0(w)$ exactly as above. But because of the fixed costs, we may now have $U^0(w) < 0$. In this case, we can reduce the expected value of output all the way to zero in the above process without having to reduce each agent's cost to zero. The worst case will then involve some agents having actions with positive costs (corresponding to having $b_i > 0$ for some agent(s) i in Lemma 5). The determination of these costs is why the solution to the minimization problem in Lemma 5 is more complicated when $U^0(w) < 0$. While there does not appear to be a simple closed-form expression for $S(w)$ in this case, the solution has the following regularity: The cost of the new action is set to zero for the agents with the smallest shares as this provides the largest reduction in the expected value of output for a given reduction in costs; the agents with the largest shares will have their costs reduced only to the minimum cost $\underline{c}_i^0$, which may be positive. We omit the details.

Second, it is worth emphasizing that even when $U^0(w) \geq 0$, the virtual surplus $U^0(w)$ is the relevant guarantee only because the true technology is unknown. Namely, any maximizer $a^* \in A^0$ in $U^0(w)$ is an equilibrium of $\Gamma(w, A^0)$ as it maximizes the potential P on A^0. (It can be verified that any such a^* is an equilibrium even if $\alpha_i = 0$ for some i, in which case P is not a potential for the game.) The resulting expected surplus is $\mathbb{E}_{F(a^*)}[v(y)] - \sum_i c_i(a_i^*)$. So if the principal knew the true technology to be A^0, he would expect (at least) this much surplus instead of $U^0(w)$. Consequently, a linear team-optimal contract is in general not an optimal linear contract for a model where the principal knows the technology.

modulo the fact that there we require incentives to be strict to ensure uniqueness of the equilibrium outcome.

1.5 Principal-Optimal Contracts

We then turn to principal-optimal contracts that maximize the principal's guaranteed expected profit $V(w)$. By Theorem 1, we can restrict attention to contracts that align the agents' interests, because any contract that fails to do so is dominated by the zero contract. However, unlike in the case of budget-balanced contracts studied in the previous section, this restriction by itself does not imply any particular relationship between the value of the outcome, $v(y)$, and the agents' compensation. Thus, even though a linear contract still turns out to be optimal, showing this requires more work than in the case of team-optimal contracts, where linearity was implied by Theorem 1.

As a first step, we derive a convenient representation of the candidate optimal contracts. Note that if w is a principal-optimal contract, then the lowest payment to each agent must be zero. Otherwise we could strictly increase the principal's profit with the contract w' defined by $w'(y) := w(y) - (\min w_1(Y), \ldots, \min w_I(Y))$ for all $y \in Y$, because subtracting the constants does not affect the agents' incentives (i.e., $\mathcal{E}(w, A) = \mathcal{E}(w', A)$ for all A), but it reduces the principal's wage bill. Moreover, because w aligns the agents' interests, there exists an outcome $\underline{y} \in Y$ that yields the zero payment simultaneously to all agents so that $w(\underline{y}) = 0$. A contract with this property is said to be *anchored at the origin.*

Any contract w that aligns the agents' interests and is anchored at the origin has the following representation. Let $\bar{w}(y) := \sum_i w_i(y)$ denote the agents' total compensation under w given outcome y. Then there exist shares $\alpha = (\alpha_1, \ldots, \alpha_I) \in [0,1]^I$, with $\sum_i \alpha_i = 1$, such that for every agent i,

$$w_i(y) = \alpha_i \bar{w}(y) \quad \forall y \in Y. \tag{1.5.1}$$

That is, each agent is paid some fixed share of the total compensation for any outcome. Conversely, any contract that can be written in this form and where $\bar{w}(y) = 0$ for some y aligns the agents' interests and is anchored at the origin. This result is almost immediate from the definitions; it can be proven the same way as Lemma 1.[9]

With the above representation in hand, we can think of the problem of finding a principal-

[9]Replace v with $\bar{w}$ everywhere in the proof of Lemma 1, and observe that $\min \bar{w}(Y) = 0$ because w is anchored at the origin.

optimal contract in two stages. First, given the agents' shares α, determine how the total compensation $\bar{w}(y)$ should depend on the outcome y. Second, optimize over the agents' shares. As the total compensation is one dimensional, the first stage resembles the single-agent case studied by Carroll (2015). We will make use of this connection to show that, here too, it is optimal to tie total compensation linearly to the value of the outcome, i.e., $\bar{w}(y) = \beta v(y)$ for some $\beta \in [0, 1]$. The payment to each agent is then $w_i(y) = \alpha_i \beta v(y)$, implying that a linear contract is optimal. (Recall that a contract is said to be linear if each agent is paid a fixed share of the value $v(y)$ for each y; the shares may or may not sum to one across agents.) The second stage then consists of finding the optimal linear contract by optimizing jointly with respect to β and α.

More precisely, we have the following result:

Theorem 4. *(i) There exists a linear principal-optimal contract.*

(ii) A principal-optimal contract w guarantees a positive expected profit (i.e., $V(w) > 0$) if and only if the known technology A^0 satisfies (1.4.8).

(iii) If every action profile in the known technology A^0 satisfies full support, then every principal-optimal contract that guarantees a positive expected profit is linear.

Theorem 4 parallels the findings for team-optimal contracts in Theorem 2. Parts (i) and (iii) extend Carroll's (2015) result on the optimality of linear contracts to multiple agents. Heuristically, a linear contract aligns interests not just among the agents, but among all the parties. This prevents the agents from seeking personal gain at the other agents' or the principal's expense.

Part (ii) gives a necessary and sufficient condition for the profit guarantee to be positive. This condition is the same one that characterizes whether a team-optimal contract can guarantee positive expected surplus. To see why this is the case, note that the principal can ensure a positive profit by choosing β close enough to 1 and taking α to coincide with the shares in the team-optimal contract whenever the latter generates a positive expected surplus. In contrast, if the condition fails, then the free-rider problem makes the surplus guarantee negative for every linear contract with $\beta = 1$; decreasing the agents' shares by

lowering β can only make matters worse. Note that with one agent, (1.4.8) reduces to the requirement that some known action generates a positive expected surplus, which is precisely Carroll's (2015) non-triviality assumption.

The rest of this section is devoted to the proof of Theorem 4. Along the way, we obtain a formula for the profit guarantee for a linear contract, which can be used to find the agents' shares in a linear principal-optimal contract. We remark on this after the proof.

We will make use of the fact that—as far as the agents are concerned—any contract w that aligns the agents' interests and is anchored at the origin can be interpreted as a budget-balanced contract in an auxiliary model where the value of each outcome y is given by $\bar{w}(y)$. This allows us to recycle results, most notably the characterization behind Lemma 5, from the analysis of team-optimal contracts. (In case of the zero contract $w \equiv 0$, the auxiliary model does not satisfy the non-triviality assumption $\max v(Y) = \max \bar{w}(Y) > 0$, so we treat it separately.) This allows the analysis here to proceed relatively fast.

The virtual surplus $U^0(w)$ continues to play a key role in the analysis. To avoid confusion, we write

$$\bar{U}^0(w) := \sup_{a \in A^0} \left(\mathbb{E}_{F(a)}[\bar{w}(y)] - \sum_i \frac{c_i(a_i)}{\alpha_i} \right) \in \mathbb{R} \cup \{-\infty\},$$

where we have simply replaced $v(y)$ with the total compensation $\bar{w}(y)$ in the definition of $U^0(w)$ in (1.4.2), and where $0/0 = 0$ and $x/0 = \infty$ for $x > 0$ by convention.

To shorten the statements of some of the lemmas that follow, we say that a contract w is *eligible* if it (i) aligns the agents' interests, (ii) is anchored at the origin, and (iii) satisfies $V(w) > 0$ and $V(w) \geq V(0)$. This definition adapts Carroll's (2015) notion of an eligible contract to the multi-agent setting, parts (i) and (ii) being the novel requirements. For example, any linear contract satisfies (i) and (ii). But the set of eligible contracts may nevertheless be empty, since the best profit guarantee may be zero in violation of (iii). However, if $V(w) > 0$ for some contract w, then this contract is eligible unless $V(w) < V(0)$, in which case the zero contract is eligible. In particular, any principal-optimal contract with a positive guarantee is eligible.

The following characterization is analogous to the single-agent case.

Lemma 8. *Let w be an eligible contract, different from the zero contract. Then*

$$V(w) = \min_{G \in \Delta(Y)} \mathbb{E}_G[v(y) - \bar{w}(y)] \quad \textit{subject to} \quad \mathbb{E}_G[\bar{w}(y)] \geq \bar{U}^0(w). \tag{1.5.2}$$

Moreover, if G achieves the minimum, then $\mathbb{E}_G[\bar{w}(y)] = \bar{U}^0(w)$.

The proof of Lemma 8 relies on the characterization behind Lemma 5. To see that $V(w)$ is not less than the minimum, interpret w as a budget-balanced contract in a model where $v(y) = \bar{w}(y)$. Then Lemma 6 implies that every technology $A \supseteq A^0$ has an equilibrium a^* such that $\mathbb{E}_{F(a^*)}[\bar{w}(y)] \geq \mathbb{E}_{F(a^*)}[\bar{w}(y)] - \sum c_i(a_i^*)/\alpha_i \geq \bar{U}^0(w)$. Thus the principal's profit is at least the minimum profit under distributions satisfying the constraint in (1.5.2).

We prove the other direction in the Appendix. The key observation is that $(E, b) \in \mathbb{R}_+^{I+1}$ defined by $E = \mathbb{E}_G[\bar{w}(y)]$ and $b = 0$ is a feasible point in the minimization problem in Lemma 5. Thus, the constructive direction in the proof of Lemma 5 gives us a technology $A \supseteq A^0$ where the unique equilibrium distribution of outcomes is approximately G.

An application of Lemma 8 yields a formula for the profit guarantee for any eligible linear contract w, where $\bar{w}(y) = \beta v(y)$ for some $\beta \in (0, 1]$ and $w_i(y) = \beta\alpha_i v(y)$. It turns out that the formula is also valid for the zero contract whenever it is eligible.[10]

Lemma 9. *Let w be an eligible linear contract with $\beta \in [0, 1]$ and $\alpha \in [0, 1]^I$, $\sum \alpha_i = 1$. Then*

$$V(w) = (1 - \beta) \max_{a \in A^0} \left(\mathbb{E}_{F(a)}[v(y)] - \sum_i \frac{c_i(a_i)}{\beta\alpha_i} \right), \tag{1.5.3}$$

where $0/0 = 0$ and $x/0 = \infty$ for $x > 0$ by convention.

Note that Lemma 9 applies also when $I = 1$, in which case $\alpha_1 = 1$ and (1.5.3) reduces to the formula derived for the single-agent case by Chassang (2013) and Carroll (2015).

Proof. If w is different from the zero contract and G achieves the minimum in (1.5.2), then

$$V(w) = (1 - \beta)\mathbb{E}_G[v(y)] = \frac{1 - \beta}{\beta}\mathbb{E}_G[\bar{w}(y)] = \frac{1 - \beta}{\beta}\bar{U}^0(w).$$

[10]The zero contract has a continuum of parameterizations by α and β, since any α_i with $\sum_i \alpha_i = 1$ will do if $\beta = 0$. This multiplicity creates no problem in the analysis, but of course it could be avoided by denoting the agents' shares by $\gamma_i := \beta\alpha_i$, in which case $\beta = \sum \gamma_i \leq 1$ and the zero contract corresponds to $\gamma = 0$.

The claim now follows by writing out $\bar{U}^0(w)$ and noting that the supremum is achieved, since Lemma 8 implies that $\bar{U}^0(w) \geq 0$.

If the zero contract is eligible, then there exists a profile $a \in A^0$ such that $c(a) = 0$ and $\mathbb{E}_{F(a)}[v(y)] > 0$. (We can't have $\underline{c}_i^0 > 0$ for any i, because $V(0)$ could then be driven to zero by giving agent i a cheaper action—cf. footnote 6.) Any $a \in A^0$ with $c(a) = 0$ is an equilibrium given any technology $A \supseteq A^0$, and hence $V(0) = \max \mathbb{E}_{F(a)}[v(y)]$ over $a \in A^0$ such that $c(a) = 0$. This agrees with the formula in the lemma, given the conventions involving 0. □

From (1.5.3) we can deduce the existence of a best linear contract and see that (1.4.8) is necessary and sufficient for the profit that it guarantees to be positive.

Lemma 10. *There exists a linear contract w^* such that $V(w^*) \geq V(w)$ for every linear contract w. Moreover, $V(w^*) > 0$ if and only if the known technology satisfies* (1.4.8).

Proof. If no linear contract is eligible, then the zero contract has $V(0) = 0$ and thus it is optimal within the class of linear contracts. If there exists an eligible linear contract, then the claim follows by continuity of (1.5.3) in β and α_i.

The derivation leading to Theorem 3 in Section 1.4 shows that if (1.4.8) holds, then we have $U^0(w) = \max_{a \in A^0}(\mathbb{E}_{F(a)}[v(y)] - \sum c_i(a_i)/\alpha_i) > 0$ for some $\alpha_i \in [0,1]^I$ with $\sum \alpha_i = 1$. Thus, (1.5.3) is positive for β close enough to 1. Conversely, if (1.4.8) does not hold, then we have $(1-\beta)\max_{a \in A^0}(\mathbb{E}_{F(a)}[v(y)] - \sum c_i(a_i)/\beta\alpha_i) \leq (1-\beta)U^0(w) \leq 0$ for every β and α, showing that (1.5.3) is nonpositive. Hence, no linear contract is eligible. □

With these facts regarding linear contracts established, to prove Theorem 4 it suffices to show that any eligible contract can be improved upon by a linear contract, strictly so if every known action profile satisfies full support. To this end, fix an eligible contract and consider the representation (1.5.1). We will show that the contract can be (weakly) improved upon by making the total compensation $\bar{w}$ a linear function of the output value $v(y)$, while keeping each agent's share α_i of the total compensation fixed. As the total compensation is one-dimensional, this allows us to draw on the proof of the single-agent case.

We need the following key lemma from the single-agent case. It identifies a particular supporting hyperplane to the set of pairs $(v(y), \bar{w}(y))$ under contract w that will be used to define the improvement contract.

Lemma 11. *Let w be an eligible contract, different from the zero contract. Then there exist numbers κ and λ, with $\lambda > 0$, such that*

$$v(y) - \bar{w}(y) \geq \kappa + \lambda \bar{w}(y) \quad \forall y \in Y, \tag{1.5.4}$$

$$V(w) = \kappa + \lambda \bar{U}^0(w). \tag{1.5.5}$$

Proof. The claim follows from Lemma 3 in Carroll (2015).[11] □

Given an eligible contract w and numbers κ, λ satisfying (1.5.4) and (1.5.5), define the contract w' by

$$\bar{w}'(y) := \frac{1}{1+\lambda} v(y) - \frac{\kappa}{1+\lambda} \quad \text{and} \quad w_i'(y) := \alpha_i \bar{w}'(y), \tag{1.5.6}$$

where α_i is agent i's share in the representation (1.5.1) of the original contract w. Then $\bar{w}'(y) \geq \bar{w}(y) \geq 0$ for all $y \in Y$ by (1.5.4), and thus $w_i' \geq 0$ for all i as required by our definition of a contract. Note that $\bar{w}'(y_0) = -\kappa/(1+\lambda) \geq 0$ implies $\kappa \leq 0$.

The affine contract w' so defined will be shown to improve on the contract w. Moreover, w' can be further improved upon simply by removing the constant payment, which does not affect the agents' incentives, and which is nonnegative by limited liability. Because w' is affine, it is not anchored at the origin and hence it neither has the representation (1.5.1), nor is it eligible. So for technical reasons, it is convenient to show both improvements at once.

To this end, define the linear contract w'' by setting

$$w_i''(y) := \frac{\alpha_i}{1+\lambda} v(y) = w'(y_i) + \frac{\alpha_i \kappa}{1+\lambda} \leq w_i'(y), \tag{1.5.7}$$

where the inequality holds because, as noted above, $\kappa \leq 0$.

Lemma 12. *Let w be an eligible contract, different from the zero contract, that satisfies (1.5.4) and (1.5.5), and let w'' be the linear contract defined by (1.5.7). Then $V(w'') \geq V(w)$. Moreover, if every known action profile satisfies full support and w is not linear, then $V(w'') > V(w)$.*

[11] The only point where a clarification is warranted is when arguing that the numbers λ and μ—the latter only appears in the proof—are positive. Then Carroll (2015) shows that $V(w) < V(0)$ (or $V_P(w) < V_P(0)$ in his notation), contradicting the eligibility of w. He cites the proof of his Lemma 1. For us the corresponding reference is the proof of our Lemma 8, where this same conclusion is arrived at by appealing to our Lemma 4.

Proof. We observe first that since $\bar{w}(y) \leq \bar{w}'(y) = \bar{w}''(y) - \kappa/(1+\lambda)$ for all $y \in Y$, we have

$$\bar{U}^0(w) \leq \max_{a \in A^0} \left(\mathbb{E}_{F(a)} \left[\bar{w}''(y) - \frac{\kappa}{1+\lambda} \right] - \sum_i \frac{c_i(a_i)}{\alpha_i} \right) = \bar{U}^0(w'') - \frac{\kappa}{1+\lambda}. \tag{1.5.8}$$

The contract w'', being linear, satisfies (1.5.1). Reinterpreting it as a budget-balanced contract we apply Lemma 6 (with the substitutions $v(y) = \bar{w}''(y)$ and $U^0(w'') = \bar{U}^0(w'')$) to find for any $A \supseteq A^0$ a pure-strategy equilibrium $a^* \in \mathcal{E}(w'', A)$ with $\mathbb{E}_{F(a^*)}[\bar{w}''(y)] \geq \bar{U}^0(w'')$. But $\mathcal{E}(w', A) = \mathcal{E}(w'', A)$ as the constants do not affect incentives. Thus, $a^* \in \mathcal{E}(w', A)$ and

$$\mathbb{E}_{F(a^*)}[\bar{w}'(y)] = \mathbb{E}_{F(a^*)}[\bar{w}''(y)] - \frac{\kappa}{1+\lambda} \geq \bar{U}^0(w'') - \frac{\kappa}{1+\lambda} \geq \bar{U}^0(w),$$

where the last inequality is by (1.5.8). Moreover, w' satisfies (1.5.4) by construction, and thus

$$V(w', A) \geq \mathbb{E}_{F(a^*)}[v(y) - \bar{w}'(y)] \geq \kappa + \lambda \mathbb{E}_{F(a^*)}[\bar{w}'(y)] \geq \kappa + \lambda \bar{U}^0(w) = V(w), \tag{1.5.9}$$

where the last step is by (1.5.5). Because A was arbitrary, this implies $V(w') \geq V(w)$. Now $V(w'') = V(w') - \kappa/(1+\lambda) \geq V(w')$ shows that $V(w'') \geq V(w)$ as desired.

It remains to show the strict inequality for non-linear contracts under the full support assumption. Observe first that if w' is not linear (i.e., if $\kappa < 0$), then $V(w'') > V(w') \geq V(w)$. So suppose w' is linear. Let every action profile in A^0 satisfy full support, i.e., $F(a) \neq \delta_{y_0}$ implies $\operatorname{supp} F(a) = Y$ for all $a \in A^0$. If w is not linear, then $\bar{w}(y) \leq \bar{w}'(y) = \bar{w}''(y) - \kappa/(1+\lambda)$ holds with strict inequality for some $y \in Y$. Furthermore, because w is eligible, we have $\bar{U}^0(w) > 0$ by Lemma 8, and so the maximum in $\bar{U}^0(w)$ is achieved by some $a \in A^0$ such that $F(a)$ has full support. This implies that the inequality in (1.5.8) is strict. The strict inequality carries through to imply that in (1.5.9), $V(w', A)$ is bounded above $V(w)$ uniformly in $A \supseteq A^0$. Therefore, $V(w'') \geq V(w') > V(w)$. □

We can now summarize how the claims in Theorem 4 follow from the previous lemmas.

Proof of Theorem 4. For part (i), we may restrict attention to contracts that align the agents' interests and are anchored at the origin. Let w be any such contract. If w is not eligible, then

it is dominated by the zero contract, which is linear. If w is eligible, then Lemmas 11 and 12 imply that there exists a linear contract that does at least as well as w. Thus, either way, w is weakly dominated by a linear contract, and so the existence of a linear principal-optimal contract follows from Lemma 10, which shows the existence of a best linear contract.

Part (ii) follows from part (i) and Lemma 10.

For part (iii), suppose that every action profile in A^0 satisfies full support. Let w be a nonlinear principal-optimal contract with $V(w) > 0$. Then w is eligible and it can be strictly improved upon with a linear contract by Lemmas 11 and 12, a contradiction. □

A linear principal-optimal contract can be found as follows. If (1.4.8) is not satisfied, then the zero contract is optimal. Otherwise, maximize (1.5.3) with respect to β and α.

In the latter case, similarly to team-optimal contracts, we may first find the optimal shares for each $a \in A^0$, and then maximize with respect to a. So fix $a \in A^0$. Note that given any β, maximizing (1.5.3) with respect to α gives us the team-optimal shares in an auxiliary model where the value of each outcome is $\beta v(y)$. These are given by $\alpha(a)$ defined in (1.4.7). Similarly, maximizing (1.5.3) with respect to β given α gives us the principal-optimal share in an auxiliary single-agent model where each action $a \in A^0$ costs the agent $C(a, \alpha) := \sum_i c_i(a_i)/\alpha_i$. Carroll (2015) shows that this is given by $\beta(a, \alpha) = \sqrt{C(a, \alpha)/\mathbb{E}_{F(a)}[v(y)]}$. Substituting these shares back into the objective gives

$$\left(\sqrt{E_{F(a)}[v(y)]} - \sqrt{C(a, \alpha(a))}\right)^2 = \left(\sqrt{E_{F(a)}[v(y)]} - \sqrt{\sum_{i,j} \sqrt{c_i(a_i)c_j(a_j)}}\right)^2.$$

Maximizing the above expression (or the square root thereof to eliminate the square) yields a maximizer a^*, which can be substituted back into the formulas for the shares to obtain the optimal contract $\alpha^* = \alpha(a^*)$ and $\beta^* = \beta(a^*, \alpha^*)$.

The simplest non-trivial examples can be created by assuming that there is only one known action profile so that $A^0 = \{a^0\}$. In that case there is no maximization over actions, so the optimal linear contract is given by $\alpha^* = \alpha(a^0)$ and $\beta^* = \beta(a^0, \alpha^*)$.

1.6 Concluding Remarks

We have shown that demanding team incentives to be robust to nonquantifiable uncertainty about the game played by the agents leads to contracts that align the agents' interests. Such contracts have a natural interpretation as being team-based. Under budget balance they reduce to linear contracts, showing that profit-sharing, or equity, is a team-optimal contract. And the contract with the best profit guarantee for the principal is similarly linear.

These incentive schemes have two additional robustness properties, which play no role in our analysis, but likely contribute to their popularity. First, interest alignment limits the scope for collusion among subsets of agents as all agents' payoffs are already similar to each other, save for the costs. Moreover, in case of a linear principal-optimal contract, the agents' compensation varies linearly with the value of the outcome, so any collusive scheme that increases the agents' total compensation also increases the principal's payoff.

Second, interest alignment not only limits a contract's downside, it also has the potential to increase the upside as it gives the agents an incentive to take advantage of unexpected opportunities to help each other and to allocate tasks efficiently. This upside potential is lost on our worst-case analysis, and it is unclear how to capture it short of moving to a fully Bayesian framework. (See Itoh (1991) and Garicano and Santos (2004) for the analysis of incentives to help and to allocate tasks in Bayesian models.) Note, however, that if the said opportunities are not unexpected (i.e., if they are part of the known technology), then they do affect our analysis: withholding help or not referring a task to a better-equipped agent are examples of the kind of negative actions that lead to the worst case.

The driving force behind our results is that only by completely eliminating conflict in the agents' preferences over outcomes can the contract guarantee good performance in all games. Taking the worst case over all games consistent with the known technology is arguably a strong assumptions. On one hand, it leads to a tractable analysis and yields sharp predictions about optimal contracts. But on the other hand, the prediction that all teams be governed via linear schemes is obviously empirically false.

With this in mind, one may view the contribution of this paper to be that it identifies robustness as a force pushing towards contracts that align the agents' interests. The analysis

here shows that if this concern is strong enough, only linear schemes survive as optimal ones. Robustness is, however, only one of many considerations affecting contract design. Consequently, the incentive schemes we observe in practice reflect it to varying degrees. A natural way to try to incorporate this into the worst-case analysis would be to restrict the set of games deemed possible, with smaller sets then resulting in less limitations on contract form. Identifying subsets of games for which the analysis remains tractable is a nontrivial problem which we leave for future work.

1.7 Appendix

1.8.1 Proofs for Section 1.3

Proof of Lemma 2. We assume throughout the proof that the set Y^* is empty.

We use the following notation. Let $\bar{w}_i := \max_{y \in Y} w_i(y)$ and $Y_i^* := \arg\max_{y \in Y} w_i(y)$. The projection of W to the payments of agents i and j is denoted $W_{i,j} \subset \mathbb{R}^2_+$. The interior of $W_{i,j}$ relative to $\mathbb{R}^2$ is $\text{int}(W_{i,j})$. For any $x \in W$, we write $x_{i,j}$ for the image of x in $W_{i,j}$.

To simplify the exposition, we assume that w_i is not constant for any agent i. We comment at the end of the proof how the argument needs to be adjusted to accomodate such agents. (These agents can be essentially ignored when constructing the worst-case technology, so the issue is mostly notational.)

A Preliminary Technology. We construct a technology with a unique equilibrium expected payment profile in W. It forms the basis of all other technologies used in the proof.

Define A^1 by letting $A_i^1 = A_i^0 \cup \{a_i^1\}$ for all i. Let $c_i(a_i^1) = \underline{c}_i^0$ for all i so that the new action is a least-cost action for each agent.

For each i, fix $z^i \in W$ such that $z_i^i = \bar{w}_i$. Let $F^i \in \Delta(Y)$ be such that $\mathbb{E}_{F^i}[w(y)] = z^i$.

Given an action profile a in A^1 such that at least one agent plays the new action a_i^1, let $n = n(a) := |\{i : a_i = a_i^1\}|$. We then define the outcome distribution by setting

$$F(a) = \left(1 - \frac{I - n + 1}{I}\xi\right)\frac{1}{n}\sum_{i\,:\,a_i = a_i^1} F^i + \frac{I - n + 1}{I}\xi H,$$

where $\xi \in (0,1)$, and H is the uniform distribution on Y. The corresponding profile of expected payments to the agents is

$$x(a) = \left(1 - \frac{I-n+1}{I}\xi\right)\frac{1}{n}\sum_{i:\, a_i = a_i^1} z^i + \frac{I-n+1}{I}\xi\mathbb{E}_H[w(y)]. \tag{1.8.1}$$

We record the following observations for future reference:

1. Since $w_i(y)$ is not constant in y, a_i^1 is the unique best-response to any profile a_{-i} where some agent $j \neq i$ plays a_j^1. This is because playing a_i^1 shifts the convex combination in (1.8.1) in the direction of z^i and reduces the weight on the full-support distribution H; the latter effect gives uniqueness even if we have $z_i^i = z_i^j$ for all j such that $a_j = a_j^1$.

2. The distribution $F(a)$ has full support on Y. Hence, $x_i(a) < \bar{w}_i$ for every agent for whom w_i is not constant. Moreover, $x_{i,j}(a)$ is in the interior of the projection $W_{i,j}$ (relative to $\mathbb{R}^2$) for all agents i and j for which the interior is nonempty.

Lemma 1.8.1. *The profile a^1 is the unique equilibrium of $\Gamma(w, A^1)$ for all $\xi > 0$ small enough.*

Proof. Observation 1. implies that a^1 is an equilibrium, and that it is the only equilibrium where at least one agent i plays the new action a_i^1. Thus it only remains to show that this is the case in every equilibrium.

Let $E := \cap_i E_i$, where $E_i := \{a \in A^1 : a_i \in A_i^0\}$. Then E is the event that every agent plays some known action. Suppose towards contradiction that there exists an equilibrium σ such that $\sigma(E) = \prod_i \sigma_i(A_i^0) > 0$. Because Y^* is empty, there exists some agent j such that, conditional on E, $F(\sigma)$ assigns probability at most $(I-1)/I$ to $\arg\max_{y \in Y} w_j(y)$. Thus, agent j's payoff, given E, is at most $\frac{I-1}{I}\bar{w}_i + \frac{1}{I}\max\{w_i(y) : y \notin Y_i^*\}$. In particular, some $\hat{a}_j \in A_j^0$ with $\sigma_j(\hat{a}_j) > 0$ yields agent j a payoff no greater than this conditional on $\cap_{i \neq j} E_i$.

Note that agent j's payoff from a_j^1 is $(1-\xi)z_j^j + \xi\mathbb{E}_H[w_j(y)] = (1-\xi)\bar{w}_j + \xi\mathbb{E}_H[w_j(y)]$ conditional on $\cap_{i \neq j} E_i$. Moreover, a_j^1 gives a strictly higher payoff than $\hat{a}_j$ if some agent $i \neq j$ plays $a_i \notin A_i^0$ by observation 1. Thus, for $\xi > 0$ small enough, a_j^1 yields a strictly higher (unconditional) expected payoff than $\hat{a}_j$, contradicting $\sigma_j(\hat{a}_j) > 0$. The cutoff for ξ depends on j, but not on σ, so it can be chosen uniformly as there are finitely many agents. □

In what follows, we assume that ξ is small enough for the result to apply.

The Worst-case, Case 1. There are two cases to consider. We first deal with the easier case where $w_i(y_0) \geq x_i(a^1)$ for some agent i, where $x(a^1)$ is the expected payment profile defined by equation (1.8.1). For concreteness, suppose the inequality holds for agent 1. We will ensure strict incentives by using the following perturbation. Let $F_\varepsilon \in \Delta(Y)$ be such that $F_\varepsilon(y_0) > 1 - \varepsilon$ and $\mathbb{E}_{F_\varepsilon}[w_1(y)] > w_1(y_0)$. Such a distribution can be found because $x_1(a^1) < \bar{w}_1$ by observation 2. F_ε will be our equilibrium distribution, so letting $\varepsilon \to 0$ will then give the desired sequence of technologies as $F_\varepsilon \to \delta_{y_0}$.

Define A by setting $A_1 = A_1^1 \cup \{a_1^2\}$, with $c_1(a_1^2) = \underline{c}_1^0$, and $A_i = A_i^1$ for $i > 1$, where A^1 is the technology constructed above. Let $F(a) = F_\varepsilon$ for all $a \in A$ such that $a_1 = a_1^2$.

We claim that F_ε is the unique equilibrium outcome distribution for the game $\Gamma(w, A)$. Indeed, the profile $(a_1^2, a_2^1, \ldots, a_I^1)$ is an equilibrium, since $\mathbb{E}_{F_\varepsilon}[w_1(y)] > x_1(a^1) > x_1(a_1, a_{-i}^1)$ for all $a_1 \in A_1^0$, and a_i^1 is optimal for agents $i > 1$ as their actions do not affect the outcome when $a_1 = a_1^2$. Moreover, agent 1 must play a_1^2 in any equilibrium where $\sigma_i(A_i^0) = 0$ for some agent i. To see this, suppose to the contrary that some such equilibrium has $\sigma_1(a_1^2) < 1$. Since $\sigma_i(A_i^0) = 0$ for some i, observation 1. then implies that a_j^1 strictly dominates all $a_j \in A_j^0$ for every agent j. Thus, $\sigma_j(A_j^0) = 0$ for all j, and hence $a_j = a_j^1$ for all $j > 1$. But then a_1^2 is agent 1's unique best-response, contradicting $\sigma_1(a_1^2) < 1$.

It remains to show that we have $\sigma_i(A_i^0) = 0$ for at least one agent i in every equilibrium of $\Gamma(w, A)$. This follows by the same argument as Lemma 1.8.1. Define the events E_i analogously and suppose there exists an equilibrium σ with $\sigma(E) > 0$. The second paragraph in the proof of Lemma 1.8.1 applies verbatim. The only difference is in the third paragraph. Now conditional on some agent $i \neq j$ playing $a_i \notin A_i^0$, a_j^1 may be only weakly better than $\hat{a}_j$: if $j \neq 1$ and $a_1 = a_1^2$, then agent j's action doesn't affect the outcome, but a_j^1 is still optimal as it is a least-cost action. This is enough to get the contradiction, because a_j^1 is strictly better than $\hat{a}_j$ conditional on $\cap_{i \neq j} E_i$. This completes the proof for the first case.

The Worst-case, Case 2. The more challenging case obtains if $w_i(y_0) < x_i(a^1)$ for all i. Then some projection $W_{i,j}$ of W has a nonempty interior relative to $\mathbb{R}^2$. To see this,

note that if $\operatorname{int} W_{i,j}$ is empty for all pairs i, j, then each $W_{i,j}$ is a (possibly degenerate) line segment. But Y^* is empty, so some line segment $W_{i,j}$ must be strictly decreasing, implying that $w_k(y_0) \geq x_k(a^1)$ for $k = i$ or $k = j$. Relabeling if necessary, we assume that $\operatorname{int} W_{1,2} \neq \emptyset$.

Consider a technology A where $A_i = A_i^1 \cup \{a_i^2, \ldots, a_i^K\}$ for $i = 1, 2$, with K a number to be specified, and where $A_i = A_i^1$ for $i > 2$. We let $c_i(a_i^k) = \underline{c}_i^0$ for all i and k so that any action $a_i \notin A_i^0$ is a least-cost action for agent i.

We first define a collection of points used to define expected payments to action profiles containing actions in $A_i \setminus A_i^1$, $i = 1, 2$. Fix $\varepsilon > 0$ and $F_\varepsilon \in \Delta(Y)$ such that $F_\varepsilon(y_0) > 1 - \varepsilon$ and $\mathbb{E}_{F_\varepsilon}[w_{1,2}(y)] \in \operatorname{int}(W_{1,2})$. F_ε will be our equilibrium outcome distribution. (We can simply take $F_\varepsilon = \delta_{y_0}$, if $w_{1,2}(y_0) \in \operatorname{int}(W_{1,2})$.) Fix $\underline{x} \in W$ such that $\mathbb{E}_{F_\varepsilon}[w_{1,2}(y)] > \underline{x}_{1,2}$.

Let x^0 be some point in $X := \{x(a) : a \in A^1,\ a \notin A^0\}$ (with $x(a)$ defined by (1.8.1)) that maximizes agent 1's payoff on X. As both $x_{1,2}^0$ and $\mathbb{E}_{F_\varepsilon}[w_{1,2}(y)]$ are in the interior of the convex set $W_{1,2}$ (the former by observation 2.), we can choose from W points $x^1, \ldots, x^{2(K-1)}$, with $x^{2(K-1)} = \mathbb{E}_{F_\varepsilon}[w(y)]$, such that the sequence $(x^0, \ldots, x^{2(K-1)})$ satisfies the following conditions (where any cases involving $k < 0$ or $k > 2(K-1)$ can be ignored):

1. Among two consecutive points, agent 1 prefers the odd one: for all $k = 0, \ldots, K-1$,

$$x_1^{2k-1} > x_1^{2k} \quad \text{and} \quad x_1^{2k+1} > x_1^{2k}.$$

2. Agent 2 has the opposite preference: for all $k = 0, \ldots, K-1$,

$$x_2^{2k-1} < x_2^{2k} \quad \text{and} \quad x_2^{2k+1} < x_2^{2k}.$$

3. Both prefer each point in the sequence to $\underline{x}$: for all $k = 0, \ldots, 2(K-1)$, $x_{1,2}^k > \underline{x}_{1,2}$.

(See Figure 1-3, where this sequence is $(x^0, x^1, x^2, x^3, x^4)$, so $K = 3$.) Note that the sequence can be constructed by first choosing a desired sequence in $W_{1,2}$, and then defining the remaining coordinates arbitrarily subject to feasibility.

With K fixed by the above sequence, we choose another sequence $(u^0, \ldots, u^{K-1})$ in W such that $u_{1,2}^{K-1} > \cdots > u_{1,2}^0$ and $x_{1,2}^k > u_{1,2}^l$ for any k and l. Say, as $x_{1,2}^k > \underline{x}_{1,2}$ for each k, we can use convex combinations of $x^{2(K-1)}$ and $\underline{x}$ with large enough weights on $\underline{x}$. (In

	A_2^0	a_2^1	a_2^2	a_2^3	$\cdots$	a_2^{K-1}	a_2^K
A_1^0	$\cdots$	$\cdot$	u^0	u^0	$\cdots$	u^0	u^0
a_1^1	$\cdot$	$\cdot$	u^1	u^1	$\cdots$	u^1	u^1
a_1^2	u^0	$\mathbf{x^1}$	$\mathbf{x^2}$	u^2	$\cdots$	u^2	u^2
a_1^3	u^0	u^1	$\mathbf{x^3}$	$\mathbf{x^4}$	$\cdots$	u^3	u^3
$\vdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\ddots$	$\cdots$	$\cdots$
a_1^{K-1}	u^0	u^1	u^2	u^3	$\cdots$	$\mathbf{x^{2K-4}}$	u^{K-1}
a_1^K	u^0	u^1	u^2	u^3	$\cdots$	$\mathbf{x^{2K-3}}$	$\mathbf{x^{2(K-1)}}$

Figure 1-5. Expected payments to agents 1 and 2 (for any fixed $a_{-\{1,2\}} \in A_{-\{1,2\}}$) in the game $\Gamma(w, A)$ in the proofs of Lemmas 2 and 3

Figure 1-3 we can take $\underline{x} = w(y_0)$, so these points would lie between x^4 and $w(y_0)$, close to $w(y_0)$.)

To complete the description of A, we assume that any action profile involving actions $\{a_i^2, \ldots, a_i^K\}$, $i = 1, 2$, leads to an expected payment profile as specified in Figure 1-5. Furthermore, we assume that the profile $x^{2(K-1)} = \mathbb{E}_{F_\varepsilon}[w(y)]$ is generated by the distribution F_ε, i.e., $F(a_1^K, a_2^K, a_{-\{1,2\}}) = F_\varepsilon$ for all $a_{-\{1,2\}} \in A_{-\{1,2\}}$. For all other expected payment profiles, any distribution $F \in \Delta(Y)$ that generates them will do.

Uniqueness of the Equilibrium Outcome Distribution in Case 2. We claim that every equilibrium of $\Gamma(w, A)$ has agents 1 and 2 playing actions a_1^K and a_2^K. Note that this leads to the outcome distribution F_ε that puts at least probability $1 - \varepsilon$ on y_0. Therefore, letting $\varepsilon \to 0$ yields a sequence of technologies with the properties listed in the lemma.

The claim follows from the following two lemmas:

Lemma 1.8.2. *Let $\sigma \in \mathcal{E}(w, A)$. If $\sigma_i(A_i^0) = 0$ for some i, then $\sigma_1(a_1^K) = 1$ and $\sigma_2(a_2^K) = 1$.*

Proof. Let $\sigma \in \mathcal{E}(w, A)$. Suppose first that $\sigma_2(A_2^0) = 0$. (The case $\sigma_1(A_1^0) = 0$ is handled analogously.) We can then eliminate column A_2^0 in Figure 1-5. Then a_1^1 strictly dominates any $a_1 \in A_1^0$ for agent 1. To see this, note that $u_1^1 > u_1^0$ by construction. And if agent 2 plays a_2^1, then a_1^1 is the unique best-response among actions in $A_1^1 = A_1^0 \cup \{a_1^1\}$ by observation 1. Therefore, $\sigma_1(A_1^0) = 0$, and we can eliminate row A_1^0 in Figure 1-5. But now that row A_1^0 and column A_2^0 are both eliminated, the remaining matrix is by construction solvable by iterated

elimination of strictly dominated strategies: a_1^2 dominates a_1^1 for agent 1 since his payoff in the cell (a_1^1, a_2^1) in Figure 1-5 is at most x_1^0, and we have $x_1^0 < x_1^1$, $u_1^1 < x_1^2$, and $u_1^1 < u_1^2$. Similarly, once a_1^1 is eliminated, a_2^2 dominates a_2^1 for agent 2 since $x_2^1 < x_2^2$, $u_2^1 < x_2^3$, and $u_2^1 < u_2^2$. Continuing iteratively, we see that only the cell (a_1^K, a_2^K) remains. (Note that we need not consider costs here as each new action is equally costly.)

Suppose then that $\sigma_i(A_i^0) = 0$ for some $i > 2$. Then the payoffs in the four cells in the top-left corner of the matrix in Figure 1-5 are given by equation (1.8.1). This implies that a_2^1 strictly dominates all $a_2 \in A_2^0$ for agent 2 by observation 1., and the fact that $u_2^1 > u_2^0$ by construction. Therefore, $\sigma_2(A_2^0) = 0$, implying that we are back in the first case. □

Lemma 1.8.3. *In every equilibrium σ of $\Gamma(w, A)$, we have $\sigma_i(A_i^0) = 0$ for some agent i.*

The proof is essentially the same as that of Lemma 1.8.1.

Proof. Define the events $E := \cap_i E_i$, $E_i := \{a \in A : a_i \in A^0\}$, and suppose $\sigma(E) > 0$ for some $\sigma \in \mathcal{E}(w, A)$. The second paragraph in the proof of Lemma 1.8.1 applies verbatim. The only difference is in the third paragraph:

If $j > 2$, then conditional on some agent $i \neq j$ playing $a_i \notin A_i^0$, a_j^1 may now be only weakly better than $\hat{a}_j$. This is because, if $a_i = a_i^k$ for some $i \in \{1, 2\}$ and $k \geq 2$, then agent j's action doesn't affect the outcome. However, a_j^1 is still a best-response as it is a least-cost action. This is enough to get the contradiction, because a_j^1 does strictly better than $\hat{a}_j$ conditional on $\cap_{i \neq j} E_i$.

If $j \in \{1, 2\}$, then a_j^1 does strictly better than $\hat{a}_j$ whenever some agent $i \neq j$ plays $a_i \notin A_i^0$ by observation 1., and the fact that $u_j^1 > u_j^0$ by construction. A contradiction. □

As promised, we comment here on how to accommodate agents for whom w_i is constant. Let J be the set of such agents. We have $|J| \leq I - 2$, since w fails to align the agents' interests. We can then apply the above construction to agents $\{1, \ldots, I\} \setminus J$, setting $A_j = A_j^0$ for all $j \in J$ and letting the new actions of agents not in J dictate the outcome in all of the technologies considered. The above analysis still applies, with the obvious modification that any "all i"-statement now means "all $i \notin J$" where relevant.[12] In every equilibrium, any

[12]In particular, Case 1 is now defined as having $w_i(y_0) \geq x_i(a^1_{-\{J\}})$ for some agent $i \notin J$, whereas in Case 2, we have $w_i(y_0) < x_i(a^1_{-\{J\}})$ for all $i \notin J$. (The notation $a^1_{-\{J\}}$ reflects the fact that $x(a)$ is still defined

agent $j \in J$ will then play some least-cost action(s) in A_j^0, without affecting the outcome. For example, in Lemma 1.8.1, the result is now that every equilibrium has each agent $i \notin J$ playing a_i^1. All other results hold as stated. □

Proof of Lemma 3. The proof has many elements in common with that of Lemma 2, but here we also need to manipulate the agents' costs, leading to some important differences.

Since $Y^* \neq \emptyset$, there exists $z \in W$ such that $z \geq x$ for all $x \in W$. (Just take $z = w(y)$ for any $y \in Y^*$.) This implies that the projection $W_{i,j}$ of W is two-dimensional for some i and j as otherwise w would align the agents' interests. Without loss, take this pair to consist of agents 1 and 2. Note that the interior of $W_{1,2}$ relative to $\mathbb{R}^2_+$, or $\text{int}(W_{1,2})$, is nonempty. Recall that $x_{1,2}$ denotes the image of $x \in W$ in $W_{1,2}$.

We will construct a technology A where $A_i = A_i^0 \cup \{a_i^1, \ldots, a_i^K\}$ for $i = 1, 2$, with K to be specified, and where $A_i = A_i^0 \cup \{a_i^1\}$ for $i > 2$.

Fix $\eta \in (0, 1)$. Let $c_i(a_i^k) = \eta \underline{c}_i^0$ for each agent i and $k = 1, \ldots, K$. Then the new actions are strictly cheaper than any known action with a positive cost: if $c_i(a_i) > 0$ for some $a_i \in A_i^0$, then at least one inequality in $c_i(a_i) \geq \underline{c}_i^0 \geq \eta \underline{c}_i^0$ is strict.

In order to define the expected payments, fix $\varepsilon > 0$ and let $F_\varepsilon \in \Delta(Y)$ be a distribution such that $F_\varepsilon(y_0) > 1 - \varepsilon$ and $\mathbb{E}_{F_\varepsilon}[w_{1,2}(y)] \in \text{int}(W_{1,2})$. (If $w_{1,2}(y_0) \in \text{int}(W_{1,2})$, we can simply take $F_\varepsilon = \delta_{y_0}$.) F_ε will be our equilibrium outcome distribution. Because $\mathbb{E}_{F_\varepsilon}[w_{1,2}(y)] < z_{1,2}$, we can then find a point $x^0 \in W$ such that (i) $z_{1,2} > x_{1,2}^0 > \mathbb{E}_{F_\varepsilon}[w_{1,2}(y)]$, (ii) $x_{1,2}^0 \in \text{int}(W_{1,2})$, and (iii) x^0 is close enough to z so that

$$\sum_i (x_i^0 - \eta \underline{c}_i^0) > \sum_i \mathbb{E}_{F(a)}[w_i(y) - c_i(a_i)] \quad \text{for all } a \in A^0. \tag{1.8.2}$$

To see why (1.8.2) can be satisfied, fix $a \in A^0$. If $\text{supp}\, F(a) \subseteq Y^*$ so that $\mathbb{E}_{F(a)}[w_i(y)] = z_i$ for all i, then by the assumption in the lemma, $c_j(a_j) > 0$ for some j. Then $\eta \underline{c}_j^0 < c_j(a_j)$ and thus the inequality in (1.8.2) holds for any x^0 such that $\sum_i x_i^0$ is close enough to $\sum_i z_i$. If $\text{supp}\, F(a) \nsubseteq Y^*$, then $\sum_i z_i > \sum_i \mathbb{E}_{F(a)}[w_i(y)]$, and again the inequality holds for any x^0 such that $\sum_i x_i^0$ is close enough to $\sum_i z_i$. As A^0 is finite, some $x^0 < z$ thus satisfies (1.8.2).

We let the profile of expected payments to be $x(a) = x^0$ for any $a \in A$ such that $a_i \in A_i^0$

by (1.8.1) to only depend on actions of agents not in J.)

for $i = 1, 2$ and $a_j = a_j^1$ for at least one agent $j > 2$.

We then define expected payments for action profiles involving new actions of agent 1 and 2. By the above conditions (i) and (ii), we can pick z^1 and z^2 in W such that $z_i^i > \mathbb{E}_{F_\varepsilon}[w_i(y)]$, $z_i^i > x_i^0 > z_i^j$, $j \neq i$. (See Figure 1-4.) Then we complete the construction of the sequence $(x^0, \ldots, x^{2(K-1)})$ and choose the sequence $(u^0, \ldots, u^{K-1})$ exactly as in the proof of Lemma 2. Finally, given any $a_{-\{1,2\}} \in A_{-\{1,2\}}$, we assign the expected payment profiles to agent 1 and 2's actions as in Figure 1-5, where the top-left corner is now given by the matrix

	A_2^0	a_2^1
A_1^0	$\cdots$	z^2
a_1^1	z^1	x^0

that reflects elements specific to the current construction. Any distributions generating these payoffs will do, except for $x^{2(K-1)}$, which is generated by the distribution $F_\varepsilon \in \Delta(Y)$.

We claim that every equilibrium of the game $\Gamma(w, A)$ so constructed has agents 1 and 2 playing a_1^K and a_2^K, and thus the unique equilibrium distribution is F_ε, which assigns at least probability $1 - \varepsilon$ to y_0. As $\min c_i(A_i) = \eta \underline{c}_i^0$, letting $\varepsilon \to 0$ and $\eta \to 1$ simultaneously (say, put $1 - \eta = \varepsilon \to 0$) yields a sequence of technologies with the desired properties.

To prove the claim, we show first that agents 1 and 2 must play a_1^K and a_2^K with probability 1 in every equilibrium σ where $\sigma_j(A_j^0) = 0$ for some agent j. Indeed, suppose this holds for some $j > 2$. Then the payoff in the cell (A_1^0, A_2^0) in Figure 1-5 is x^0. But then a_i^1 dominates all actions in A_i^0 for agents $i = 1, 2$, and iterated elimination leads to the profile (a_1^K, a_2^K) as desired. If instead $j \in \{1, 2\}$, say, $j = 1$, then the top row in Figure 1-5 is eliminated, and so a_2^1 dominates all actions in A_2^0 for agent 2, and iterated elimination again leads to (a_1^K, a_2^K). The case $j = 2$ is handled similarly.

It remains to show that every $\sigma \in \mathcal{E}(w, A)$ has $\sigma_j(A_j^0)$ for some j. Let $E := \cap_i E_i$, with $E_i := \{a \in A : a \in A^0\}$. Suppose towards contradiction that $\sigma(E) > 0$. Then (1.8.2) implies that conditional on E, the expected payoff of some agent j is strictly less than $x_j^0 - \eta \underline{c}_j^0$. Thus, some $\hat{a}_j \in A_j^0$ with $\sigma_j(\hat{a}_j) > 0$ yields agent j a payoff strictly less than $x_j^0 - \eta \underline{c}_j^0$ conditional on $\cap_{i \neq j} E_i$. Hence, a_j^1 gives a strictly higher payoff than $\hat{a}_j$ given $\cap_{i \neq j} E_i$. Moreover, if any

agent $i \neq j$ plays $a_i \notin A_i^0$, then a_j^1 still gives at least as high a payoff as $\hat{a}_j$: for $j > 2$ this is because agent j's action then doesn't affect the outcome and a_j^1 is a least-cost action; for $j = 1, 2$ this is because a_j^1 then dominates any $a_j \in A_j^0$ by construction. Therefore, a_j^1 yields a higher (unconditional) expected payoff than a_j, contradicting $\sigma_j(\hat{a}_j) > 0$. □

Proof of Corollary 1. By Lemmas 2 and 3, it suffices to show that if there exists $a^* \in A^0$ such that $\operatorname{supp} F(a^*) \subseteq Y^*$ and $c(a^*) = 0$, then there exists a technology $A \supseteq A^0$ such that for all $\sigma \in \mathcal{E}(w, A)$, we have $F(\{y \in Y : v(y) = 0\}|\sigma) = 1$. So fix any such a^*.

Note that if a^* satisfies full support, then $F(a^*) = \delta_{y_0}$ (since $Y^* \neq Y$ if w fails to align the agents' interests), and hence $y_0 \in Y^*$. If a^* satisfies costly production, then $\mathbb{E}_{F(a^*)}[v(y)] = 0$, and again $\{y \in Y : v(y) = 0\} \cap Y^* \neq \emptyset$. Thus, either way, $v(y^*) = 0$ for some $y^* \in Y^*$.

Let A be the technology constructed in the first part of the proof of Lemma 4 with $v(y^*) = 0$. As noted there, if $\sigma \in \mathcal{E}(w, A)$, then the agents only play zero cost actions under σ and $\operatorname{supp} F(\sigma) \subseteq Y^*$, since otherwise some agent i could profitably deviate to a_i'. Therefore, $\sigma(a) > 0$ for $a \in A^0$ only if (i) $c(a) = 0$ and (ii) $\operatorname{supp} F(a) \subseteq Y^*$. But we saw above that $F(\{y \in Y : v(y) = 0\}|a) = 1$ for all such a. On the other hand, if $a \notin A^0$, then $F(a) = \delta_{y^*}$, where $v(y^*) = 0$. Thus A has the desired property. □

Proof of Corollary 3. If w is budget balanced so that $\sum_i w_i(y) = v(y)$ for all $y \in Y$, then $Y^* \subseteq \arg\max_{y \in Y} \sum_i w_i(y) = \arg\max_{y \in Y} v(y)$. The assumption in Corollary 3 then implies that the case covered by Lemma 4 cannot arise, so the claim follows by Lemmas 2 and 3. □

1.8.2 Proofs for Section 1.4

We first restate and prove part (i) of Theorem 2:

Lemma 1.8.4. *There exists a linear team-optimal contract.*

Proof. By Lemma 1, we can identify the space of linear budget-balanced contracts with the compact set $B := \{\alpha \in [0,1]^I : \sum_i \alpha_i = 1\}$. We will denote such a contract simply by α. As noted in the discussion following Theorem 2 in the main text, it suffices to show that the guaranteed expected surplus $S(\alpha)$ is an upper semi-continuous function of α on B.[13]

[13]The argument that follows parallels Carroll's (2015) proof of existence of an optimal linear contract with general cost lower bounds in the single-agent case.

Fix a sequence (α^n) in B converging to some $\alpha \in B$. (Since B is finite dimensional, any norm will do.) We need to show that $S(\alpha) \geq \limsup_n S(\alpha^n)$. By moving to a subsequence if necessary, we can assume that $S(\alpha^n)$ converges to $\limsup_n S(\alpha^n)$. Fix any technology $A \supseteq A^0$ and denote by σ^n the equilibrium of $\Gamma(\alpha, A)$ that achieves $S(\alpha^n, A)$. Extracting a further subsequence if necessary, we can assume that the sequence (σ^n) converges to some $\sigma \in \Delta(A)$. Since the agents' payoffs are continuous in α, the profile σ is an equilibrium of $\Gamma(\alpha, A)$ by the upper hemi-continuity of the Nash equilibrium correspondence. We thus have

$$\begin{aligned} S(\alpha, A) &\geq \mathbb{E}_{F(\sigma)}[v(y)] - \sum_a \sigma(a) \sum_i c_i(a) \\ &= \lim_n \left(\mathbb{E}_{F(\sigma^n)}[v(y)] - \sum_a \sigma^n(a) \sum_i c_i(a) \right) = \lim_n S(\alpha^n, A) \geq \lim_n S(\alpha^n). \end{aligned}$$

Since $A \supseteq A^0$ was arbitrary, this implies $S(\alpha) \geq \lim_n S(\alpha^n)$ as desired. □

We then prove Lemma 6.

Proof of Lemma 6. Fix any w and A as in the statement of the Lemma. Define positive "auxiliary shares" $\tilde{\alpha}_i \in (0, 1]$, $i = 1, \ldots, n$, by setting $\tilde{\alpha}_i = \alpha_i$ if $\alpha_i > 0$, and otherwise letting $\tilde{\alpha}_i > 0$ be any number small enough such that

$$\tilde{\alpha}_i < \frac{\min\{|c_i(a_i) - c_i(a_i')| : a_i, a_i' \in A_i,\ c_i(a_i) \neq c_i(a_i')\}}{\max v(Y)}.$$

(Note that $\sum_i \tilde{\alpha}_i > 1$ if $\alpha_i = 0$ for some i.) Define the function $\tilde{P} : A \to \mathbb{R}$ by

$$\tilde{P}(a) := \mathbb{E}_{F(a)}[v(y)] - \sum_i \frac{c_i(a_i)}{\tilde{\alpha}_i}. \tag{1.8.3}$$

It is straightforward to verify that then, for every agent i and every a_i, a_i', and a_{-i},

$$u_i(a_i, a_{-i}) - u_i(a_i', a_{-i}) > 0 \quad \text{implies} \quad \tilde{P}(a_i, a_{-i}) - \tilde{P}(a_i', a_{-i}) > 0.$$

Indeed, if $\alpha_i > 0$, then $u_i(a_i, a_{-i}) - u_i(a_i', a_{-i}) = \alpha_i(\tilde{P}(a_i, a_{-i}) - \tilde{P}(a_i', a_{-i}))$. If $\alpha_i = 0$, then

$u_i(a_i, a_{-i}) - u_i(a_i', a_{-i}) > 0$ implies $c_i(a_i) < c_i(a_i')$, and so the choice of $\tilde{\alpha}_i$ implies

$$\tilde{P}(a_i, a_{-i}) - \tilde{P}(a_i', a_{-i}) = \mathbb{E}_{F(a_i, a_{-i})}[v(y)] - \mathbb{E}_{F(a_i', a_{-i})}[v(y)] - \frac{c_i(a_i) - c_i(a_i')}{\tilde{\alpha}_i} > 0.$$

Thus, $\tilde{P}$ is a generalized ordinal potential for $\Gamma(w, A)$, and hence $\arg\max_{a \in A} \tilde{P}(a) \subseteq \mathcal{E}(w, A)$. In particular, there exists a pure-strategy equilibrium.

It remains to establish the inequalities. Fix an equilibrium $a^* \in \arg\max_{a \in A} \tilde{P}(a)$. Then $\alpha_i \mathbb{E}_{F(a^*)}[v(y)] - c_i(a_i^*) \geq -\underline{c}_i^0$, as otherwise agent i could deviate to a least-cost action in A_i^0. Consider then the first inequality. If $U^0(w) = -\infty$, it holds vacuously. So let $U^0(w) > -\infty$. As noted after the definition of $U^0(w)$ in (1.4.2), then $\alpha_j = 0$ implies $\underline{c}_j^0 = 0$. Moreover, we have $c_j(a_j^*) = 0$ for any such agent j, since otherwise he could deviate to a zero-cost action in A_j^0. But then $c_i(a_i^*)/\tilde{\alpha}_i = c_i(a_i^*)/\alpha_i$ for all i (since $\tilde{\alpha}_i = \alpha_i > 0$ or $c_i(a_i^*) = 0$), and we have

$$\mathbb{E}_{F(a^*)}[v(y)] - \sum_i \frac{c_i(a_i^*)}{\alpha_i} = \mathbb{E}_{F(a^*)}[v(y)] - \sum_i \frac{c_i(a_i^*)}{\tilde{\alpha}_i} = \max_{a \in A} \tilde{P}(a) \geq \max_{a \in A^0} \tilde{P}(a) \geq U^0(w),$$

where the first inequality follows since $A \supseteq A^0$, and the second follows since $\alpha_i \leq \tilde{\alpha}_i$. □

As preparation for the proof of Lemma 5, the following lemma establishes some properties of the solutions to the minimization problem.

Lemma 1.8.5. *If (E, b) achieves the minimum in Lemma 5, then*

(i) $\alpha_i = 0$ *implies* $b_i = 0$ $(\forall i = 1, \ldots, I)$.

(ii) (1.4.3) *holds with equality.*

Furthermore, there exists a minimizer (E, b) satisfying the following additional properties:

(iii) $Eb_i = 0$ $(\forall i = 1, \ldots, I)$.

(iv) $E = \max\{U^0(w), 0\}$.

(v) $b_i \leq \underline{c}_i^0$ $(\forall i = 1, \ldots, I)$.

(The proof shows that properties (iii)–(v) are in fact necessary unless $\alpha_i = 1$ for some i.)

Proof. A minimizer exists since we are minimizing a continuous function over a compact set.

We first deal with the case $U^0(w) = \max v(Y)$. It is straightforward to verify that then $(E, b) = (U^0(w), 0)$ is the only feasible point, and that it satisfies properties (i)–(v).

From now on, let $\max v(Y) > U^0(w) > -\infty$. Note that if $\alpha_i = 0$, then $\underline{c}_i^0 = 0$ (since $U^0(w) > -\infty$), and so (1.4.4) implies that only $b_i = 0$ is feasible. This shows property (i).

To show (ii), suppose to the contrary that $E - \sum_i b_i/\alpha_i > U^0(w)$ for some minimizer (E, b). Then equality must hold in (1.4.4) for all i, since otherwise we could increase some b_i. Thus, $b_i = \alpha_i E + \underline{c}_i^0$, which we can substitute for b_i in (1.4.3). Rearranging then gives

$$E(1 - |\{i : \alpha_i > 0\}|) > U^0(w) + \sum_i \frac{\underline{c}_i^0}{\alpha_i} = \max_{a \in A^0} \left(\mathbb{E}_{F(a)}[v(y)] - \sum_i \frac{c_i(a_i) - \underline{c}_i^0}{\alpha_i} \right) \geq 0,$$

where the equality is by definition of $U^0(w)$, and the last inequality follows because $a \in A^0$ with $c_i(a_i) = \underline{c}_i^0$ for all i is feasible. But $|\{i : \alpha_i > 0\}| \geq 1$, and thus $E(1 - \{i : \alpha_i > 0\}|) \leq 0$, contradicting the strict inequality in the other direction above. This establishes (ii).

If there exists some agent i with $\alpha_i = 1$, then we can clearly satisfy properties (iii)-(v) by letting $(E, b) = (U^0(w), 0)$ if $U^0(w) \geq 0$, and letting $(E, b_i, b_{-i}) = (0, -U^0(w), 0)$ otherwise.[14]

For the rest of the proof, we assume that $\alpha_i < 1$ for all i, and we show that any minimizer satisfies properties (iii)–(v). Let $b_j = 0$ for all $j \in J^0 := \{i : \alpha_i = 0\}$. Consider minimization only over E and b_i, $i \in \{1, \ldots, I\} \setminus J^0$. By inspection, the feasible set is a convex polyhedron in $\mathbb{R}^{I+1-|J^0|}$ with a nonempty interior, and the objective function is affine. The following

[14]To see that (1.4.4) and property (v) are satisfied for agent i, note that

$$\alpha_i E - b_i = E - b_i = U^0(w) = \max_{a \in A^0 : c_j(a_j) = 0 \, \forall j \neq i} \left(\mathbb{E}_{F(a)}[v(y)] - c_i(a_i) \right) \geq -\underline{c}_i^0.$$

Kuhn-Tucker conditions are thus necessary for a minimum:

$$\beta \geq 0,\ \mu_i \geq 0,\ \eta \geq 0,\ \lambda \geq 0,\ \theta_i \geq 0, \tag{1.8.4}$$

$$1 - \lambda - \sum_i \theta_i \alpha_i - \eta + \beta = 0, \tag{1.8.5}$$

$$-1 + \frac{\lambda}{\alpha_i} - \mu_i + \theta_i = 0, \tag{1.8.6}$$

$$\beta(\bar{V} - E) = 0,\ \mu_i b_i = 0,\ \eta E = 0,\ \lambda\Big(E - \sum_i \frac{b_i}{\alpha_i} - U^0(w)\Big) = 0,\ \theta_i(\alpha_i E - b_i + \underline{c}_i^0) = 0, \tag{1.8.7}$$

where i ranges over all agents with $\alpha_i > 0$, and where $\bar{V} := \max v(Y)$.

We can now show property (iii). Let (E, b) be a minimizer. If $E = 0$, then we are done, so let $E > 0$. We show first that $\theta_i = 0$ for all i. Suppose to the contrary that $\theta_i > 0$ for some i. Note that if $\theta_i > 0$, then $b_i = \alpha_i E + \underline{c}_i^0 > 0$ by (1.8.7), implying that $\mu_i = 0$. Thus, multiplying both sides of (1.8.6) by α_i and then summing over all i such that $\theta_i > 0$ gives

$$0 = \sum_{i:\theta_i>0} (-\alpha_i + \lambda) + \sum_i \theta_i \alpha_i = \sum_{i:\theta_i>0} (-\alpha_i + \lambda) + 1 - \lambda + \beta,$$

where the second equality substitutes for $\sum_i \theta_i \alpha_i$ using (1.8.5), noting that $\eta = 0$ by (1.8.7). Rearranging the terms yields

$$\sum_{i:\theta_i>0} \alpha_i = 1 + \beta + (|\{i : \theta_i > 0\}| - 1)\lambda \geq 1 + \beta \geq 1,$$

since $\theta_i > 0$ for at least one agent by assumption. But $\sum_{i:\theta_i>0} \alpha_i \geq 1$ implies that we must have $\theta_i > 0$ for all i such that $\alpha_i > 0$, and hence (1.4.4) holds as equality for every such agent by (1.8.7). Since (1.4.4) is an equality also for each agent in J^0, it thus holds as an equality for every agent, leading to a contradiction with (1.4.3) as in the proof of property (i) above. We conclude that $\theta_i = 0$ for all i. This implies that $\lambda = 1 + \beta$ by (1.8.5), so that (1.8.6) becomes

$$-1 + \frac{1 + \beta}{\alpha_i} - \mu_i = 0,$$

which in turn implies $\mu_i > 0$, as $\alpha_i < 1$ by assumption. Hence, $b_i = 0$ by (1.8.7), as desired.

Property (iv) follows from properties (ii) and (iii). Namely, if $U^0(w) \geq 0$, then only $(E, b) = (U^0(w), 0)$ is consistent with both (ii) and (iii). On the other hand, if $U^0(w) < 0$, then $b_i > 0$ for some i by (ii), which by (iii) implies $E = 0$.

It remains to show property (v). If $U^0(w) \geq 0$, then (ii) and (iv) imply $b_i = 0 \leq \underline{c}_i^0$ for all i. If $U^0(w) < 0$, then $E = 0$ by (iv), and (1.4.4) implies $b_i \leq \underline{c}_i^0$ for all i. □

The next lemma will be used to show that $S(w)$ is not greater than the minimum in Lemma 5. It will also be used to characterize $V(w)$ in Section 1.5, which is why we state it in a form that emphasizes the uniqueness of the equilibrium distribution of outcomes.

Lemma 1.8.6. *Let w be a budget-balanced contract that aligns the agents' interests. Suppose $\max v(Y) > U^0(w) > -\infty$. Let $(E, b) \in [0, \max v(Y)] \times \mathbb{R}_+^I$ and $G \in \Delta(Y)$ be such that*

(i) $E > \max\{U^0(w), 0\}$,

(ii) (1.4.3) *holds with strict inequality,*

(iii) $b_i \leq \underline{c}_i^0$ *for all i (and thus* (1.4.4) *is satisfied), and*

(iv) $\mathbb{E}_G[v(y)] = E$.

Then there exists a technology $A \supseteq A^0$ such that every $\sigma \in \mathcal{E}(w, A)$ satisfies $F(\sigma) = G$ (and hence $\mathbb{E}_{F(\sigma)}[v(y)] = E$) and $\sum_{a_i} \sigma_i(a_i) c_i(a_i) = b_i$ for all i.

Proof of Lemma 1.8.6. Fix (E, b) and G as in the lemma. Let $J^0 = \{i : \alpha_i = 0\}$. Note that $0 \leq |J^0| < I$, because w is budget balanced. Moreover, we have $\underline{c}_i^0 = 0 = b_i$ for all $i \in J^0$, since $U^0(w) > -\infty$. We will construct a technology $A \supseteq A^0$ where $A_i = A_i^0 \cup \{a_i'\}$ with $c_i(a_i') = b_i$ for all $i \notin J^0$, and $A_i = A_i^0$ for all $i \in J^0$. Equilibria of $\Gamma(w, A)$ will consist of profiles where each agent $i \notin J^0$ plays a_i' and agents in J^0 mix over zero-cost actions.

We define outcome distributions for action profiles involving the new actions as follows. Let $\varepsilon_I := E - \max\{U^0(w), 0\}$ and let $0 \equiv \varepsilon_{|J^0|} < \varepsilon_{|J^0|+1} < \cdots < \varepsilon_I$, to be used to provide strict incentives. We will assume that each ε_k is small enough to satisfy the finitely many restrictions imposed on it by (1.8.9) below. Fix $a \in A$ and let $J := \{i : a_i = a_i'\} \cup J^0$.

Suppose $|J^0| < |J|$ so that at least one agent plays the new action in a. If equality holds in[15]

$$\max_{a_J \in A_J^0} \left(\mathbb{E}_{F(a_J, a_{-J})}[v(y)] - \sum_{j \in J} \frac{c_j(a_j) - b_j}{\alpha_j} \right) \leq \max v(Y), \tag{1.8.8}$$

then we take $F(a)$ to be any distribution such that $\mathbb{E}_{F(a)}[v(y)] = \max v(Y)$. Since $b_j \leq \underline{c}_j^0$ for every agent by assumption (iii), the only other possibility is that (1.8.8) holds with strict inequality instead. In that case we let $F(a)$ to be any distribution such that

$$\mathbb{E}_{F(a)}[v(y)] = \left[\max_{a_J \in A_J^0} \left(\mathbb{E}_{F(a_J, a_{-J})}[v(y)] - \sum_{j \in J} \frac{c_j(a_j) - b_j}{\alpha_j} \right) \right]^+ + \varepsilon_{|J|} < \max v(Y), \tag{1.8.9}$$

where $[r]^+ := \max\{r, 0\}$ for $r \in \mathbb{R}$. (This defines at most finitely many inequalities involving $\varepsilon_{|J|}$, because A is finite.)

Note that if $|J| = I$ so that every agent $i \notin J^0$ plays a_i' in the profile a, then the left-hand side of (1.8.8) equals $U^0(w) + \sum_i b_i/\alpha_i < E \leq \max v(Y)$, where the strict inequality is by assumption (ii). Thus $F(a)$ satisfies (1.8.9) and $\mathbb{E}_{F(a)}[v(y)] = [U^0(w)]^+ + \varepsilon_I = E$. We may thus put $F(a) = G$ for any such a.

The following lemma is the first step towards a characterization of $\mathcal{E}(w, A)$.

Lemma 1.8.7. *Let $a \in A$ be an action profile where every agent in J^0 plays a zero-cost action. Then $u_i(a_i', a_{-i}) \geq u_i(a)$ for every agent $i \notin J^0$.*

Proof. Let $i \notin J^0$ and $\hat{a}_i \in A_i^0$. Fix $a_{-i} \in A_{-i}$ such that $c_j(a_j) = 0$ for all $j \in J^0$. Then

$$u_i(\hat{a}_i, a_{-i}) \leq \max_{a_i \in A_i^0} u_i(a_i, a_{-i}) = \alpha_i \max_{a_i \in A_i^0} \left(\mathbb{E}_{F(a_i, a_{-i})}[v(y)] - \frac{c_i(a_i)}{\alpha_i} \right). \tag{1.8.10}$$

Let $N := \{j \neq i : a_j = a_j'\}$. Suppose first that the maximum in (1.8.10) is achieved by a_i such that $F(a_i, a_{-i})$ is defined by (1.8.9), with $J = N \cup J^0$. We can then write out the far

[15]Recalling that $\underline{c}_i^0 = b_i = 0$ for all $i \in J^0$ shows that the maximum on the left-hand side is finite, since we can always choose zero-cost actions for every agent with $\alpha_i = 0$.

right-hand side of (1.8.10) as[16]

$$\begin{aligned}
&\alpha_i \max_{a_i \in A_i^0} \Big(\Big[\max_{a_J \in A_J^0} \big(\mathbb{E}_{F(a_J, a_i, a_{-J\cup\{i\}})}[v(y)] - \sum_{j \in J} \frac{c_j(a_j) - b_j}{\alpha_j} \big) \Big]^+ + \varepsilon_{|J|} - \frac{c_i(a_i)}{\alpha_i} \Big) \\
&\leq \alpha_i \min \Big\{ \Big[\max_{a_{J\cup\{i\}} \in A^0_{J\cup\{i\}}} \big(\mathbb{E}_{F(a_{J\cup\{i\}}, a_{-J\cup\{i\}})}[v(y)] - \sum_{j \in J\cup\{i\}} \frac{c_j(a_j) - b_j}{\alpha_j} \big) \Big]^+ + \varepsilon_{|J|} - \frac{b_i}{\alpha_i}, \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \max v(Y) - \frac{b_i}{\alpha_i} \Big\} \\
&\leq \alpha_i \big(\mathbb{E}_{F(a_i', a_{-i})}[v(y)] - \frac{b_i}{\alpha_i} \big) = u_i(a_i', a_{-i}),
\end{aligned} \tag{1.8.11}$$

where the first inequality uses $b_i \leq \underline{c}_i^0$ and the second inequality follows by definition of $F(a_i', a_{-i})$ (with $J = N \cup J^0 \cup \{i\}$), since $\varepsilon_{|J|} < \varepsilon_{|J\cup\{i\}|}$. Thus $u_i(\hat{a}_i, a_{-i}) \leq u_i(a_i', a_{-i})$.

If instead the maximum in (1.8.10) is achieved by $\tilde{a}_i$ such that $\mathbb{E}_{F(\tilde{a}_i, a_{-i})}[v(y)] = \max v(Y)$, then (1.8.8) holds with equality. This implies that there exists some $\tilde{a}_J \in A_J^0$ such that $c_j(a_j) - b_j = 0$ for all $j \in J$ and $\mathbb{E}_{F(\tilde{a}_J, \tilde{a}_i, a_{-J\cup\{i\}})}[v(y)] = \max v(Y)$. The right-hand side of (1.8.10) now becomes

$$\begin{aligned}
\alpha_i \big(\max v(Y) - \frac{c_i(\tilde{a}_i)}{\alpha_i} \big) &= \alpha_i \big(\mathbb{E}_{F(\tilde{a}_J, \tilde{a}_i, a_{-J\cup\{i\}})}[v(y)] - \sum_{j \in J\cup\{i\}} \frac{c_j(\tilde{a}_j) - b_j}{\alpha_j} - \frac{b_i}{\alpha_i} \big) \\
&= \alpha_i \max_{a_{J\cup\{i\}} \in A^0_{J\cup\{i\}}} \big(\mathbb{E}_{F(a_{J\cup\{i\}}, a_{-J\cup\{i\}})}[v(y)] - \sum_{j \in J\cup\{i\}} \frac{c_j(a_j) - b_j}{\alpha_j} \big) - b_i \\
&\leq \alpha_i \mathbb{E}_{F(a_i', a_{-i})}[v(y)] - b_i = u_i(a_i', a_{-i}),
\end{aligned} \tag{1.8.12}$$

where the last line follows by definition of $F(a_i', a_{-i})$ (with $J = N \cup J^0 \cup \{i\}$). This shows that $u_i(\hat{a}_i, a_{-i}) \leq u_i(a_i', a_{-i})$ in this case as well. □

Lemma 1.8.7 implies that any $\sigma \in \Delta(A)$ with $\sigma_i(a_i') = 1$ for $i \notin J^0$ and $\sigma_j(a_j)c_j(a_j) = 0$ for $j \in J^0$ is an equilibrium. In any such equilibrium, only profiles $a \in A$ with $|J| = I$ arise with positive probability. Therefore, $F(\sigma) = G$, $\mathbb{E}_{F(\sigma)}[v(y)] = E$ and $\sum_{a_i} \sigma_i(a_i)c_i(a_i') = b_i$. To rule out other equilibria, we need the following result.

[16] If $N = \emptyset$, then by convention we ignore the maximization over $a_J \in A_J^0$ and the operator $[\cdot]^+$ on the first line of (1.8.11).

Lemma 1.8.8. *Let $a \in A$ be an action profile where every agent in J^0 plays a zero-cost action. If $a_{-J^0} \neq a'_{-J^0}$, then $u_i(a'_i, a_{-i}) > u_i(a_i, a_{-i})$ for some agent i.*

Proof. Fix $a \in A$ as in the lemma. Define J as above. Assume towards contradiction that $u_i(a'_i, a_{-i}) = u(a_i, a_{-i})$ for all $i \notin J$. Then, for each agent $i \notin J$, (1.8.11) or (1.8.12) holds as a chain of equalities. By definition of the distributions $F(a'_i, a_{-i})$, this is possible only if $\mathbb{E}_{F(a'_i, a_{-i})}[v(y)] = \mathbb{E}_{F(a)}[v(y)] = \max v(Y)$ and $c_i(a_i) = b_i$ for all $i \notin J$; otherwise the choice of ε_k results in at least one strict inequality in each case. But note that, by definition of $F(a)$, we have $\mathbb{E}_{F(a)}[v(y)] = \max v(Y)$ only if

$$\begin{aligned}\max v(Y) &= \max_{a_J \in A^0_J} \left(\mathbb{E}_{F(a_J, a_{-J})}[v(y)] - \sum_{j \in J} \frac{c_j(a_j) - b_j}{\alpha_j} \right) \\ &= \max_{a_J \in A^0_J} \left(\mathbb{E}_{F(a_J, a_{-J})}[v(y)] - \sum_{i=1}^{I} \frac{c_i(a_i) - b_i}{\alpha_i} \right) \leq U^0(w) + \sum_i \frac{b_i}{\alpha_i} < E,\end{aligned}$$

where the second equality follows because $c_i(a_i) = b_i$ for all $i \notin J$, the weak inequality follows because in $U^0(w)$ the maximization is over all agents' actions, and the strict inequality is because (E, b) satisfies (1.4.3) with strict inequality by assumption (ii). Thus, $E > \max v(Y)$, contradicting $E \in [0, \max v(Y)]$. □

To complete the proof, observe that the agents in J^0 can clearly only play zero-cost actions in any equilibrium. Consider $\sigma \in \Delta(A)$ where this is true. Suppose $\sigma_i(a'_i) < 1$ for some agent $i \notin J^0$. Then some profile $\hat{a}$ satisfying the assumptions of Lemma 1.8.8 arises with positive probability under σ. Let i be an agent who can profitably deviate from $\hat{a}$ to a'_i; the existence of such an agent follows by Lemma 1.8.8. If agent i deviates from σ_i to playing a'_i for sure, then his payoff increases strictly when the other agents play $\hat{a}_{-i}$ (which happens with positive probability under σ_{-i}), and it increases weakly against all other a_{-i} by Lemma 1.8.7. Thus σ is not an equilibrium. □

We are now in a position to prove Lemma 5.

Proof of Lemma 5. By Lemma 6, every technology $A \supseteq A^0$ has a pure strategy equilibrium where the expected value of output E and costs b satisfy (1.4.3) and (1.4.4). Hence $S(w)$ is

not less than the minimum of $E - \sum b_i$ over $E \in [0, \max v(Y)]$ and $b \in \mathbb{R}^I_+$ such that (1.4.3) and (1.4.4) are satisfied.

In the other direction, suppose first that $U^0(w) < \max v(Y)$. By Lemma 1.8.5.(iv) and (v), there exists a minimizer (E, b) with $E = \max\{U^0(w), 0\}$ and $b_i \leq \underline{c}_i^0$ for all i. Let $E < E' \leq \max v(Y)$. Then (E', b) together with any G such that $\mathbb{E}_G[v(y)] = E'$ satisfies the assumptions in Lemma 1.8.6, and hence there exists a technology $A \supseteq A^0$ such that $S(w) \leq S(w, A) = E' - \sum b_i$. Letting $E' \to E$ shows that $S(w)$ is not greater than the minimum.

If on the other hand $U^0(w) = \max v(Y)$, then $(E, b) = (\max v(Y), 0)$ achieves the minimum by Lemma 1.8.5.(iii) and (iv). As $S(w) \leq \max v(Y)$ by feasibility, $S(w)$ is less than the minimum in this case as well. We conclude that $S(w)$ equals the minimum.

The properties of the minimizers follow from Lemma 1.8.5. □

1.8.3 A Proof for Section 1.5

Proof of Lemma 8. That $V(w)$ is not less than the minimum is shown in the main text.

To prove the converse, note that the feasible set in (1.5.2) is compact, so the minimum is achieved at some G^*. Let $\pi := \mathbb{E}_{G^*}[v(y) - \bar{w}(y)]$. We show below that $\bar{U}^0(w) < \max \bar{w}(Y)$. Thus we can approximate G^* with a sequence (G^n) such that $E^n := \mathbb{E}_{G^n}[\bar{w}(y)] > \bar{U}^0(w)$ and $\mathbb{E}_{G^n}[v(y) - \bar{w}(y)] \to \pi$ as the objective is continuous in G.[17] Every $(E^n, 0)$ and G^n satisfy the assumptions of Lemma 1.8.6 (with the substitutions $v(y) = \bar{w}(y)$ and $U^0(w) = \bar{U}^0(w)$), and thus there exists a technology $A^n \supseteq A^0$ for which G^n is the unique equilibrium distribution of outcomes. Hence, $V(w) \leq V(w, A^n) = \mathbb{E}_{G^n}[v(y) - \bar{w}(y)] \to \pi$ as desired.

To show that $\bar{U}^0(w) < \max \bar{w}(Y)$, suppose to the contrary that $\bar{U}^0(w) = \max \bar{w}(Y)$. The definition of $\bar{U}^0(w)$ then implies that there exists a profile $a \in A^0$ such that $c(a) = 0$ and $\operatorname{supp} F(a) \subseteq \arg\max_{y \in Y} \bar{w}(y) = Y^*$, where the equality holds because w aligns the agents' interests. Thus $V(w) < V(0)$ by Lemma 4, contradicting the eligibility of w.

It remains to show that any minimizer satisfies the constraint with equality. Let G^* be a minimizer. Because w is eligible, we have $V(w) = \mathbb{E}_{G^*}[v(y) - \bar{w}(y)] > 0$. Observe that if $\mathbb{E}_{G^*}[\bar{w}(y)] > \bar{U}^0(w)$, then the mixture $G := (1-\varepsilon)G^* + \varepsilon\delta_{y_0}$ is feasible for $\varepsilon > 0$ small enough.

[17] For example, take G^n to be the mixture $(1 - \frac{1}{n})G^* + \frac{1}{n}\delta_{\bar{y}}$ for some $\bar{y} \in \arg\max_{y \in Y} \bar{w}(y)$.

But $v(y_0) - \bar{w}(y_0) = -\bar{w}(y_0) \leq 0$, implying that $\mathbb{E}_G[\bar{w}(y) - \bar{w}(y)] \leq (1-\varepsilon)V(w) < V(w)$, which contradicts G^* being a minimizer. We conclude that $\mathbb{E}_{G^*}[\bar{w}(y)] = \bar{U}^0(w)$. $\square$

Chapter 2

Dynamic Pricing and Allocation in Ride-sharing Markets

2.1 Introduction

Ride-sharing is a new industry that has been gaining popularity fast in the past several years with the development of mobile internet technology. As the ride-sharing platform in a traffic network observes real-time supply and demand, it can make pricing and dispatching decisions accordingly, which used to be impossible in the traditional taxi industry. Dynamic pricing is an important feature in popular ride-sharing platform like Uber and Lyft, and is believed to be an important source of economic efficiency by many since it tends to help with balancing supply and demand, and has been studied intensively in the literature. What is perhaps less noticed is the central platform's ability to potentially allocate supply to demand in different locations of the traffic network in anticipation of the future. Since the ride-sharing market of a city is a dynamic inter-connected system of network, a careful modeling is needed in order to answer questions about optimal pricing and allocation policy as a whole.

In this paper we formalize a model of ride-sharing markets as follows (Section 2). A city is a fully connected network of locations, and every minute there are stochastic demands from riders to move from locations to locations. For simplicity, on the supply side we assume that the platform has full control over a fleet of vehicles,[1] and every minute can make pricing and

[1]This is a big simplification for the current situation in the real world where drivers make their own

dispatching decisions based on the current demand and supply situation. The platform's objective is to maximize its long-term profit, which is formalized as the expected average profit given an infinite horizon and no discounting.

Naturally we are interested in knowing the optimal pricing and dispatching policy. Though characterization of the optimal policy is provided, there is no explicit solution, and numerically finding one is computationally costly because an optimal policy would have to specify what to do given every possible combination of car distribution in the network and demand realization, and these are extremely large discrete spaces in a reasonably sized market. Therefore we turn to find heuristic solutions that has provably good performance.

We give the first main theorem in Section 3, which provides an upper bound on the payoff of the optimal policy by resorting to a deterministic version of this problem. In the deterministic problem, demands are deterministic with values equal to the means in the stochastic system, and we look for the prices and dispatching allocation such that the system is in steady state (i.e. the number of cars leaving any location should be equal to the number of cars coming to that location) and the profit is maximized. The deterministic problem is a simple convex contrained optimization problem and is easy to solve computationally; furthermore, we can show that the maximized profit in the deterministic problem provides an upperbound for our original problem. Later on, by comparing the performance of any heuristic policy to this upper bound, we are then able to provide a lower bound on how this policy performs relative to the optimal policy.

Next, in Section 4 we propose and analyze several heuristic policies which are also inspired by the deterministic version of the problem. All the heuristics we propose here are fixed-price heuristics, which means that the pricing decision does not depend on the current state of car distribution, thus would not change over time. This may sound to be against the whole idea of "dynamic pricing"; however we can prove that, under our assumption of stationary demand, one of the fixed-price heuristic policy has pretty good performance in thick markets. We then provide some simulation evidence that verifies and demonstrates the result, and further suggests that a second heuristic policy we consider might have at least equally good

decisions about participation and whether to accept a riding request based on prices and anticipation; however it may not be that far from reality once self-driving car technology matures and becomes widely applied.

asymptotic performance, and probably strictly better performance in small markets.

In Section 5 we explore the question of how the heuristic policies compare to the optimal policy in a small market. We first propose an algorithm that can correctly solve for the optimal policy. Even though applying it to a reasonably sized market is prohibitively costly, we can nevertheless use it to find the optimal policy in a tiny market with only two locations and two cars. Based on this, we then provide some comparisons between the optimal policy and various heuristics, and investigate possible sources of inefficiency of heuristics based on the small sample economy.

We review some related literature in the next subsection. Section 6 concludes and discusses about directions for future research.

2.1.1 Literature Review

The main research area relevant to this study is dynamic pricing. Ma, Fang, and Parkes (2018) study a ride-sharing market with complete information, in which the ride-sharing platform tries to smooth pricing in space and time to incentivize drivers to accept dispatches, by adopting a spatio-temporal pricing mechanism. Castillo, Knoepfle, and Weyl (2017) find that traditional price regulated taxi industry has to set a uniform high level of prices to avoid efficiency impairing dislocation of the drivers, and the capability of setting dynamic pricing for ride-sharing platforms can reduce or avoid the driver dislocation problem and thus charge a lower price overall. Banerjee, Riquelme, and Johari (2015) study the optimal pricing strategies for a ride-sharing platform through a model based on queueing theory. Similar to what I find in this paper, they find that the performance under dynamic pricing strategies cannot exceed that under the optimal static pricing policy, although the dynamic pricing policies are more robust to model mis-specification. The model in their paper focuses a single location with stationary system state, while my model in this paper is more general in the sense that it considers a optimization problem over a network, and the ride-sharing platform has additional degrees of freedom in setting policies, as it not only sets prices, but also matches rides and relocates vehicles. Figliozzi, Mahmassani, and Jaillet (2007) discuss a vehicle routing problem in a competitive environment in which the sequential pricing problem is treated as a sequential auction. The authors propose an dynamic pricing scheme is that is

more appropriate than a first-price auction to reflect a carrier's incremental cost of servicing rides. Dynamic pricing problems have also been studied in related optimization problems such as inventory and revenue management (Rana and Oliveira (2014), Besbes and Zeevi (2009), Gallego and Van Ryzin (1997), Gallego and Van Ryzin (1994)), and online retailing auctions (Carvalho and Puterman (2005), Raju, Narahari, and Ravikumar (2003)).

Another related strand of literature is resource optimization problems over networks. Similar to my setting, Afèche, Liu, and Maglaras (2018) study the optimization problem of a ride-sharing platform in a spatial network where the ride-sharing platform can match rides and relocate vehicles. They investigate the properties of equilibrium in a relative small scale economy with a two-location, four-route loss network. Hyytiä, Penttinen, and Sulonen (2012) build a model using Markov decision processes and M/M/1 queue for a single vehicle in dynamic transportation system, and study properties of optimization objectives that result in control policies with desired performance. Yang, Jaillet, and Mahmassani (2004) study a multi-vehicle truckload pickup and delivery system and suggest optimization policies using mixed-integer programming.

2.2 Model

I formulate the optimization problem for a ride-sharing platform serving a traffic network using a dynamic programming setting with infinite horizon and discrete time. In each period, the ride-sharing platform sets prices of trips in the traffic network and relocates vehicles to maximize long-run average payoffs.

2.2.1 Setup

Let $\{1, 2, ...L\}$ $(L \geq 2)$ be a finite set of locations, and we adopt i, j, k to denote generic locations. A city is a fully connected network with these L locations. A ride-sharing firm which is assumed to be a monopolist operates in this city a fleet of vehicles that move riders around. The fleet consists of N vehicles, and a vehicle can carry only one rider at a time.

Time is discrete[2] and horizon is infinite.

At the beginning of each time period t, the demand of rides from location i to location j, denoted as ω_{ij}^t, is assumed to be a random variable of Poisson distribution whose parameter λ_{ij}^t is a function of $p_{ij}^t \in [0, p_{max}]$, the current price for an i-to-j ride, with p_{max} being some upper bound of admissible prices. The relationship between λ_{ij}^t and p_{ij}^t is assumed to be described by a time-independent function $\lambda_{ij}(\cdot)$, i.e. $\lambda_{ij}^t = \lambda_{ij}(p_{ij}^t)$. Demand realizations across time and across different origin-destination pairs are independent. For any $i \neq j$, $\lambda_{ij}(\cdot)$ is bounded and strictly decreasing, so $\lambda_{ij}(\cdot)$ has an inverse demand function $p_{ij}(\cdot)$; $\lambda_{ii}(\cdot) \equiv 0$ is assumed for all i, which has the interpretation that either there is no such demand or such request will not be accepted ("origin too close to destination"). I further assume that for $i \neq j$, the revenue rate function, defined as $r_{ij}(\lambda) \stackrel{def}{=} \lambda p_{ij}(\lambda)$, satisfies $lim_{\lambda \to 0} r_{ij}(\lambda) = 0$, is continuous, bounded and concave, and has a unique bounded maximizer $\lambda^* = argmax_{\lambda \geq 0} r(\lambda)$.

The demand in the network is called *regular* if all the above assumptions are satisfied. In Section 4 and 5 where numeric examples are involved, I in particular assume that $\lambda_{ij}(p_{ij})$ takes the functional form of $\lambda_{ij}(p_{ij}) = \Lambda_{ij} e^{-\alpha_{ij} p_{ij}}$, where $\Lambda_{ij} \geq 0$ and $\alpha_{ij} \geq 0$ are parameters.[3] It can be verified that such demands are indeed regular.

The ride-sharing platform makes multiple choices in each period to maximize its average payoff in each period in the long-run. The ride-sharing platform is able to charge different prices for different origin-destination rides, and different prices for the same origin-destination rides at different times. In addition to the pricing decision, the ride-sharing platform also makes two other decisions:[4]

[2]The length of a discrete period is assumed to be exogenously given. For example, because of some higher-level constraint, it is just that given the pricing, matching and relocation decisions are updated every 30 seconds. I make this assumption because potentially the length of a period can be a decision variable: if we assume that requests must be replied within a period, then a longer period may correspond to a more "thick" market - for example, more requests will be accumulated in 60 seconds as opposed to 10 seconds, so if the matching decision is to be made with a larger "batch", the matching is probably more efficient, at the expense of longer waiting time.

[3]An interpretation of these parameters can be that λ_{ij} is the arrival rate of potential riders, whose willingness-to-pay has exponential distribution with parameter α_{ij} (so the average willingness-to-pay is $\frac{1}{\alpha_{ij}}$). It can be seen that if the i-to-j market is a static, isolated market, then a monopolist in this market with zero marginal cost should set the price as $\frac{1}{\alpha_{ij}}$ to maximize profit/revenue.

[4]This paper assumes that the ride-sharing platform controls the fleet, so as to simplify the analysis. In reality, it is more common to see the ride-sharing platform making the "matching" decision only and leave the drivers to relocate themselves. However, we can also interpret the relocation decision of the ride-sharing

1. Matching/dispatching. When requests arrive and multiple drivers are available, there are potentially many ways to match passengers and drivers. The ride-sharing platform assigns each rider to a driver.

2. Relocation. When a vehicle becomes idle and is not assigned any rider, the ride-sharing platform decides whether the idle vehicle should stay at the current location or should relocate to a different location.

I assume the following restrictions on matching[5]: 1. Prompt response, a ride request has to be responded in a concurrent period or rejected; 2. No car-pooling, a rider can only be matched with an idle driver; 3. Same-location match, a rider must be matched with a driver at the same location.

I assume that the travel time between two locations $t_{ij} \geq 1$ to be deterministic multiples of periods[6]. I make a trivial assumption $t_{ii} = 1$ to simplify state notations below, which simply means that a vehicle that stays at the current location would be available for matching at the same location in the next period.

With the above simplifying restrictions and assumptions, the optimization problem for the ride sharing platform can be formulated as a dynamic programming problem with infinite horizon and average-payoff criterion.

The State Space (Vehicle Locations). The state at time t for a vehicle indexed by $n \in \{1, ..., N\}$ is denoted as *vehicle state* $x_t^n = (j, \tau)$, where $j \in \{1, ..., L\}$ is the vehicle's next available-for-match (traveling destination) location and τ is the number of periods for the vehicle to arrive at location j. For example, vehicle state $(j, 0)$ means that the vehicle

platform in our setting to be giving instructions (e.g. pushing real-time traffic notifications) to drivers and direct them to desired locations. Also, some ride-sharing companies like Uber are experimentally operating a fleet of auto-driving cars in which Uber has fully control of the flow of the cars. Therefore, the assumption that the ride-sharing platform can make relocation decision is not very unrealistic.

[5]The matching decision can be very complicated in general. For example, consider the case when a request originating from i arises in period t, but currently there is no idle driver at location i; however there are several idle drivers at locations that are not too far from i. If we assume that riders are willing to wait some time, we may allow a rider at i to be matched with some idle driver at a "nearby" location. Actually, we may as well match the rider with a *busy* driver who is expected to drop off its current passenger at some nearby location soon. Or we don't even have to reply to the request immediately, but can wait several periods to see if better matching options arise later.

[6]This assumption means that it always takes the same time to travel from one place to another, ignoring traffic conditions. This is not true in real life, but it holds in a long-run equilibrium as discussed in this paper.

is currently at location j and is available for matching immediately, while $(j, 1)$ means that the vehicle is currently traveling thus unavailable for matching, but will arrive at location j and become available in one period. There are finite number of feasible vehicle states: (i, τ) is a feasible vehicle state if $0 \leq \tau \leq \max_k (t_{ki} - 1)$[7]. Let C denote the set of feasible vehicle states. The *system state* at time t is the aggregation of all vehicle states, represented by vector $x_t = (N^t(i, \tau))_{(i,\tau)\in C}$ where $N^t(i, \tau)$ denotes the number of cars in vehicle state (i, τ) at time t. $N^t(i, \tau)$ takes values between 0 and N, and naturally a feasible system state satisfies

$$\sum_{(i,\tau)\in C} N^t(i, \tau) = N. \tag{2.2.1}$$

Thus the (system) state space is also finite, denoted as X.

The Control Space (Pricing and Dispatching). At time t, a control u_t contains three decisions. The price decision $\left(p_{ij}^t\right)_{i,j\in\{1,\ldots,L\}}$[8] specifying cost of trips between any two locations, the matching decision $\left(b_{ij}^t(\omega_t)\right)$ specifying the number of vehicles that are sent from location i to j with a rider, and the relocation decision $e_{ij}^t(\omega_t)$ specifying the number of cars that are relocated from location i to j without a rider. Notice that matching decision and relocation decision constitute a complete contingent-plan specifying the matching and relocation choices for each possible situation of realized demand $\omega_t = \left(\omega_{ij}^t\right)$ (ω_{ij}^t is the number of ride requests from i to j that arise in period t).

Denote the set of feasible controls u_t when the state is x_t as $U(x_t)$. A feasible control $u_t = \left(\left(p_{ij}^t\right), \left(b_{ij}^t(\omega_t)\right), \left(e_{ij}^t(\omega_t)\right)\right)$ in system state $x_t = (N^t(i, \tau))_{(i,\tau)\in C}$ needs to satisfy the following conditions

$$\sum_{j\in\{1,\ldots,L\}} \left(b_{ij}^t(\omega_t) + e_{ij}^t(\omega_t)\right) = N^t(i, 0) \quad \forall i, \omega_t$$

$$0 \leq b_{ij}^t(\omega_t) \leq \omega_{ij}^t, \;\; e_{ij}^t(\omega_t)) \geq 0 \;\; \forall i, j, \omega_t$$

The above constraints reflect our requirements that every vehicle that is available at some

[7]Traveling time between location k and location i is t_{ki}. When a vehicle is firstly matched to a k-to-i ride, its state is $(k, 0)$, after one period, its state becomes $(i, t_{ki} - 1)$.

[8]For the rest of the paper I will simply write (p_{ij}) instead of $(p_{ij})_{i,j\in\{1,\ldots,L\}}$, with the understanding that i and j are reserved notations for locations.

location i must either be assigned a ride or get relocated (it can be relocated to the same place i), and that a car can only be matched with at most one request at a time.

The Disturbance Space (Ride Requests). The ride requests at time t, $\omega_t = \left(\omega_{ij}^t\right)$, is a random vector of realized ride requests, whose elements are independent random variables with Poisson distribution $\omega_{ij}^t \sim Poisson\left(\lambda_{ij}\left(p_{ij}^t\right)\right)$. The probability distribution of ω_t depends on the current control u_t (which specifies p_{ij}^t), but not on values of prior disturbances $\omega_{t-1}, ..., \omega_0$. Let W denote the set of ride requests.

System Dynamics. If a vehicle is in state (i, τ) where $\tau \geq 1$, then in the following period its state becomes $(i, \tau - 1)$; if a vehicle is in state $(i, 0)$ and either get matched with a i-to-j request or get relocated from i to j (one of these must happen), its state becomes $(j, t_{ij} - 1)$ in the next period. Denote the state transition function as $f(\cdot)$, which is a function of the current state, control chosen by the ride-sharing platform and realization of requests, and thus

$$x_{t+1} = f(x_t, u_t, \omega_t).$$

A vehicle in state (i, τ) can only come from two sources: either it was in state $(i, \tau + 1)$ last period, or it was in state $(k, 0)$ where k is some location that satisfies $t_{ki} - 1 = \tau$ and it was dispatched from k to i (either with or without a rider).

$$N^{t+1}(i, \tau) = N^t(i, \tau + 1) + \sum_{k:t_{ki}-1=\tau} \left(b_{ki}^t(\omega_t) + e_{ki}^t(\omega_t)\right) \tag{2.2.2}$$

Per Stage Payoffs. We assume that the cost for the ride-sharing platform to run the fleet is zero.[9] Each period, profit is generated when requests are accepted and assigned to drivers. The per period payoff is thus given by the total of cash flow received from all trips:

$$g(x_t, u_t, \omega_t) = \sum_i \sum_j p_{ij}^t b_{ij}^t(\omega_t)$$

An *admissible policy* $\pi = \{\mu_0, \mu_1, ...\}$ consists of a sequence of functions where μ_t maps

[9]This can also be interpreted as a fixed operating decision, i.e. the fleet of N vehicles has been decided to operate in the system forever, thus although in each period positive operating costs are incurred (e.g. gas cost), this cost is assumed to be independent of the matching/relocation/pricing decisions (imagine that the cars are running in the system anyway so no matter where they go, they incur a constant rate of gas cost per minute), thus can be viewed as fixed cost.

states x_t into a distribution over controls u_t in $U(x_t)$, $\mu_t\left(u_t \middle| x_t\right) \in \Delta(U(x_t))$, for all $x_t \in X$ and all $t = 0, 1, \ldots$. In other words, for period t and for all possible system states x_t, μ_t specifies the (possibly randomized) pricing decisions and allocations plans. Realizations of controls u_t in different periods are independent. If $\mu_t(x_t)$ is a degenerate distribution for all t and x_t, we say that policy π is *non-randomized*, and in this case we just conveniently write $\mu_t(x_t) = \left(\left(p_{ij}^t(x_t)\right), \left(b_{ij}^t(x_t, \omega_t)\right), \left(e_{ij}^t(x_t, \omega_t)\right)\right)$. We denote by Π the set of all (possibly randomized) admissible policies.

Given an initial state x_0 and an admissible policy $\pi = \{\mu_0, \mu_1, \ldots\}$, the expected average payoff is defined as

$$J_\pi(x_0) \stackrel{def}{=} \limsup_{T\to\infty} \frac{1}{T} E\left[\sum_{t=0}^{T-1} g(x_t, u_t, \omega_t)\right].$$

The objective of the ride-sharing platform is to maximize the expected average payoff $J_\pi(x_0)$ over all admissible policies $\pi \in \Pi$, for some given x_0.[10] The optimal average profit function $J^\star$ is defined as

$$J^\star(x) \stackrel{def}{=} \max_{\pi\in\Pi} J_\pi(x), \ x \in X. \tag{2.2.3}$$

Stationary Policies and Markov Chain Notation

A policy is said to be *stationary* if it has the form $\pi = \{\mu, \mu, \ldots\}$, meaning the control in each period is the same, which is referred to as a stationary policy and it is simply denoted as μ. For a stationary policy μ, we denote $J_\mu(x_0)$ as the expected average payoff of μ starting at x_0. As shown below, for the purpose of finding an optimal policy, it suffices to consider only stationary policies because the optimal average payoff can be attained through a stationary policy.

As the state space X is finite, we can restate the problem by introducing the following finite-state Markov Chain notations. Denoting the number of states as $S = |X|$, we can index the states as $\{1, \ldots, S\}$. Denote the probability of transiting from state x to state y

[10] A more widely seen formulation of dynamic optimization problem with infinite horizon is the expected discounted total payoff criterion. However, given the short time duration of a period in our application (about seconds or minutes per period), any reasonable discounting factor should be very close to 1 anyway, thus it would be interesting to see how things work in the "limit case" (i.e. with undiscounted payoff).

given control u as $\rho_{xy}(u)$, which is defined as:

$$\rho_{xy}(u) = P\left(x_{t+1} = y \middle| x_t = x, u_t = u\right) = P\left(W_{xy}(u) \middle| x, u\right)$$

where $W_{xy}(u)$ is defined as

$$W_{xy}(u) = \left\{\omega \in W \middle| f(x, u, \omega) = y\right\}.$$

Each time the system is in state x and control $u = (p_{ij}, b_{ij}(\omega), e_{ij}(\omega)) \in U(x)$ is applied, the expected instantaneous payoff can be written as:

$$g(x, u) \stackrel{def}{=} E[g(x, u, \omega)] = E\left[\sum_i \sum_j p_{ij} b_{ij}(\omega)\right] = \sum_i \sum_j p_{ij} E[b_{ij}(\omega)].$$

Note that $b_{ij}(\omega) \leq \omega_{ij}$ (i.e. the number of matched rides must be smaller than or equal to the number of requests), so $g(x, u) \leq \sum_i \sum_j p_{ij}(x) E\omega_{ij} = \sum_i \sum_j p_{ij}(x) \lambda_{ij}(p_{ij}(x)) = \sum_i \sum_j r_{ij}(\lambda_{ij}(p_{ij}(x)))$, which is bounded from our assumption that the revenue function $r_{ij}(\cdot)$ is bounded.

The objective of the ride-sharing platform of maximizing the average payoff per period starting from a given initial state x_0, over all policies $\pi = \{\mu_0, \mu_1 ...\}$ with $\mu_t(x) \in \Delta(U(x))$, for all x and t, can be reformulated as

$$J_\pi(x_0) = \limsup_{T \to \infty} \frac{1}{T} E\left[\sum_{t=0}^{T-1} g(x_t, \mu_t(x_t))\right]$$

where with some abuse of notation,

$$g(x_t, \mu_t(x_t)) \stackrel{def}{=} E_{u_t \sim \mu_t\left(u_t \middle| x_t\right)} g(x_t, u_t).$$

For a stationary policy μ, we denote $J_\mu(x_0)$ as the average revenue of μ starting at x_0, and we use the following shorthand notations:

$$g_\mu = \begin{pmatrix} g(1, \mu(1)) \\ \vdots \\ g(S, \mu(S)) \end{pmatrix}$$

is the vector of expected payoffs for each state s given policy μ.

$$P_\mu = \begin{pmatrix} \rho_{11}(\mu(1)) & \cdots & \rho_{1S}(\mu(1)) \\ \vdots & \ddots & \vdots \\ \rho_{S1}(\mu(S)) & \cdots & \rho_{SS}(\mu(S)) \end{pmatrix}$$

is the transition matrix of the Markov process, where each element

$$\rho_{xy}(\mu(x)) = E_{u \sim \mu\left(u \middle| x\right)} [\rho_{xy}(u)]$$

represents the expected probability of transiting from state x to state y, by sticking to the stationary policy μ.

$$J_\mu = \begin{pmatrix} J_\mu(1) \\ \vdots \\ J_\mu(S) \end{pmatrix}$$

is the vector of average payoff for each state.

Observe that

$$J_\mu = \limsup_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} P_\mu^t g_\mu,$$

which is a mathematical fact that (e.g. see Proposition 5.1.1 of Bertsekas (2005)) for any transition probability matrix P, $P^\star \stackrel{def}{=} \lim_{T\to\infty} \frac{1}{T}\Sigma_{t=0}^{T-1} P^t$ exists and satisfies $P^\star = PP^\star = P^\star P = P^\star P^\star$. Therefore it follows immediately that

$$J_\mu = P_\mu^\star g_\mu$$

where $P_\mu^\star \stackrel{def}{=} \lim_{T\to\infty} \frac{1}{T}\Sigma_{t=0}^{T-1} P_\mu^t$ and its xyth component is the long-term frequency of visits to state y given that the initial state is x.

Fix any initial state x_0. Let $\rho^T = (\rho_1, ..., \rho_S)$ be the x_0th row of the matrix $P_\mu^\star$, thus ρ is a (column) vector whose xth element is the long-term frequency of visits to state x given that the initial state is x_0. We thus have

$$J_\mu(x_0) = \rho^T g_\mu = \sum_{x \in \{1,...,S\}} \rho_x g\left(x, \mu\left(x\right)\right) \tag{2.2.4}$$

The above equation has a very intuitive interpretation: the average profit of stationary policy μ given an initial state x_0 is a weighted sum of expected instantaneous profits under all states, with the weights being the long-term frequencies under policy μ given initial state x_0.

2.2.2 Stationary Optimal Policy

In this subsection I show that for any initial state x, there exists a stationary optimal policy μ (i.e. $J_\mu\left(x\right) = J^\star\left(x\right)$).

Consider that the ride-sharing platform chooses a stationary optimal policy to maximize its long-term average payoff in each period. Similar to a dynamic programming setting of maximizing total discounted payoff, the optimization problem for the ride-sharing platform can be reformulated in a recursive fashion, namely an average-payoff version of Bellman's equation:

$$\lambda + h\left(x\right) = \max_{\mu \in \Delta(U(x))} \left[g\left(x, \mu\right) + \sum_{y=1}^{S} \rho_{xy}\left(\mu\right) h\left(y\right) \right], \ x = 1, ..., S \tag{2.2.5}$$

in which $h(x)$ and λ are unknowns are to be solved. There is a natural interpretation thatλ represents the long-term average payoff per period, and $h\left(x\right)$represents the long-term value of the ongoing business of the ride-sharing platform given system state x, which is similar to the value function in a dynamic programming setting with discounted payoffs.

The following theorem justifies existence of a solution to the above Bellman equation.

Theorem 5. *There exists a scalar λ and a real-valued function h that solve Bellman's equation (2.2.5), therefore the optimal average payoff $J^\star\left(x\right)$ does not depend on x. Furthermore, there exists a non-randomized stationary optimal policy that attains the maximum in Bell-*

man's equation (2.2.5).

To understand the above result, we first introduce the following established result in average-payoff dynamic programming theory (e.g. see Proposition 5.6.1 in Bertsekas (2005)):

Fact. *Consider a dynamic optimization problem with finite space (the control space can be finite or infinite). If a scalar λ and a real-valued function h solve Bellman's equation (2.2.5), then λ is the optimal average payoff $J^{\star}(x)$ for all $x = 1, ..., S$ (which implies that all initial states have the same optimal average payoff). Furthermore, if $\mu^{\star}(x)$ attains the maximum on the right-hand side of Bellman's equation for each x, then the stationary policy $\mu^{\star}$ is optimal, i.e. $J_{\mu^{\star}}(x) = \lambda = J^{\star}(x)$ for all x.* [11]

The above result shows that the solutions to the average payoff dynamic programming problem can be identified if we could find λ and h satisfying the average-payoff Bellman's equation (2.2.5), but it per se provides no assurance on the existence of λ and h. In fact in general such solution may not exist.[12] However, for our specific problem we can show the

[11]The proof of the result is actually straightforward and I give an outline here: Consider any admissible policy $\pi = (\mu_0, \mu_1, ...)$, and suppose λ and h satisfied Bellman's equation. Then, for any $x \in \{1, ..., S\}$ and t,

$$g(x, \mu_t(x)) + \sum_{y=1}^{S} \rho_{xy}(\mu_t(x)) h(y) \leq \lambda + h(x)$$

therefore

$$\begin{aligned} E\left[\sum_{t=0}^{T-1} g(x_t, \mu_t(x_t))\right] &\leq E\left[\sum_{t=0}^{T-1}\left(\lambda + h(x_t) - \sum_{y=1}^{S} \rho_{x_t y}(\mu_t(x)) h(y)\right)\right] \\ &= \sum_{t=0}^{T-1}\left[\lambda + Eh(x_t) - Eh(x_{t+1})\right] \\ &= T\lambda + Eh(x_0) - Eh(x_T) \end{aligned}$$

thus

$$\begin{aligned} J_{\pi}(x_0) &= \limsup_{T\to\infty} \frac{1}{T} E\left\{\sum_{t=0}^{T-1} g(x_t, \mu_t(x_t))\right\} \\ &\leq \limsup_{T\to\infty} \frac{1}{T}\left[T\lambda + Eh(x_0) - Eh(x_T)\right] \\ &= \lambda \end{aligned}$$

[12]In fact with average-payoff criterion, the general form of Bellman's equation should be a coupled pair of equations, which can accommodate the case where the optimal average payoff starting from different initial states are different thus are not characterized by a single λ (see Bertsekas (2005) Proposition 5.1.8).

existence of the solution by the following vanishing discount factor approach, which at the same time would allow us to have a better understanding of the average-payoff version of Bellman's equation.

We introduce the standard β-discounted version of the problem ($\beta < 1$) and the associated Bellman's equation

$$J_\beta^\star(x) = \max_{\mu \in \Delta(U(x))} \left[g(x,u) + \beta \sum_{y=1}^{S} \rho_{xy}(u) J_\beta^\star(y) \right]$$

Since the per period payoff is bounded, it is well known that the above β-discounted Bellman's equation has a unique bounded solution of $J_\beta^\star$ (see, for example, Proposition 1.2.3 in Bertsekas (2005)). Subtracting $\beta J_\beta^\star(x_0)$ (where x_0 is an arbitrarily picked state) from both sides of this equation and introducing the function

$$h_\beta(x) = J_\beta^\star(x) - J_\beta^\star(x_0)$$

we obtain

When the state space and the control space is finite, this coupled pair of equations always have a solution because Blackwell optimal policy always exists in this case (see Bertsekas (2005) Proposition 5.1.3) and its gain-bias pair satisfies the coupled pair of equations (see Bertsekas (2005) Proposition 5.1.4). Very typically (under some weak conditions) this coupled pair of optimality equations reduces to the single equation (2.2.5). Also, since a Blackwell optimal policy (which by definition is a stationary policy) exists and is optimal over all policies in the average payoff problem (Bertsekas (2005) Proposition 5.1.7), we know that in a finite-state, finite-control problem, a stationary optimal policy always exists.

However without finiteness of the control space or the state space (note that in this paper the control space is infinite while the state space is finite), the above results no longer hold: the coupled pair of optimality equations may not have a solution, or there may not exist an optimal stationary policy, or even the optimal average payoff may not be approachable by the payoffs of stationary policies. Quoting Bertsekas (2005), "contrary to the case of problems with finite state and control spaces, there is no comprehensive theory for infinite-spaces average cost problems".

One counterexample with finite states and infinite controls is the following: There are two states, 1 and t. At state 1, we can choose a control $u \in [0,1]$, while gaining a payoff u in this stage; we then move to state t with probability u^2, and stay in state 1 with probability $1-u^2$. State t is absorbing and has zero per-stage payoff. In this example, the coupled pair of optimality equations (which can be shown must reduce the single Bellman's equation) have no (real) solution. What is happening here is that all policies have average payoff 0, yet starting from state 1, there is an infinite advantage in total payoff relative to starting from state t. (But since all policies have zero average payoff, stationary optimal policy exists in this example.) See Bertsekas (2005) Example 3.2.1 and Example 5.6.1 for a more thorough analysis of this example.

Examples that shows non-existence of stationary optimal (or near-optimal) policy also exist, though most (if not all) of them seem to involve infinite state space. See Bertsekas (2005) Example 5.6.5 and Example 5.6.6.

$$(1-\beta)\, J_\beta^\star(x) + \beta h_\beta(x) = \max_{\mu \in \Delta(U(x))} \left[g(x,\mu) + \beta \sum_{y=1}^{S} \rho_{xy}(\mu)\, h_\beta(y) \right]. \tag{2.2.6}$$

We may view $h_\beta(x)$ as a relative payoff of state x (relative to state x_0) for the β-discounted problem. The preceding equation resembles Bellman's equation for the average payoff problem. If we take the limit of both sides as $\beta \to 1$, and suppose that the limits of all terms exist, we obtain Bellman's equation for the average payoff problem with

$$\lambda = \lim_{\beta \to 1} (1-\beta)\, J_\beta^\star(x_0),$$

$$h(x) = \lim_{\beta \to 1} h_\beta(x).$$

It turns out that in our problem, those limits indeed all exist. The crucial condition that turns out to hold under our model assumption is the following:

Lemma 13. *$|h_\beta(x)|$ is uniformly bounded over x and $\beta \in (0,1)$, that is, there exists some $B > 0$ such that $|h_\beta(x)| < B$ for all x and $\beta \in (0,1)$.*

Intuitively, Lemma 13 means that in the β-discounted problem, the difference between $J_\beta^\star(x)$ and $J_\beta^\star(x_0)$ is bounded, the optimal value achieved starting from x_0 is not very different from the optimal value achieved starting from another state x. Intuitively this should be true because starting from any system state x_0, one can always arrive at any system x within finite periods of time. See Appendix for proof details.

Let $\{\beta_k\}$ be a sequence such that $\beta_k \to 1$. Using the boundedness of h_β as well as the boundedness of $(1-\beta)\, J_\beta^\star(x_0)$ over $\beta \in (0,1)$, we can find a subsequence, also denoted $\{\beta_k\}$ for simplicity, such that $\lim_{k\to\infty} h_{\beta_k}(x) = h(x)$ for all x, and $\lim_{k\to\infty} (1-\beta_k)\, J_{\beta_k}^\star(x_0) = \lambda$, for some λ and $h(x)$. Taking the limit in equation (2.2.6) along the subsequence $\{\beta_k\}$ and interchanging limit and maximization (using the compactness of $\Delta(U(x))$ and continuity of $g(x,\mu) + \beta \sum_{y=1}^{S} \rho_{xy}(\mu)\, h_\beta(y)$ in β and μ), we just obtain the average-payoff Bellman's equation.

Based on the discussions above, the existence of a stationary optimal policy is simply a result of compactness of $\Delta(U(x))$ and continuity of $g(x,\mu) + \beta \sum_{y=1}^{S} \rho_{xy}(\mu)\, h_\beta(y)$ in μ which guarantees the existence of some $\mu^\star(x) \in \Delta(U(x))$ that attains the maximum in the

average-payoff Bellman's equation, and by the previous Fact (2.2.2), the stationary policy $\mu^\star$ is optimal. Furthermore, it is not hard to see that we can require $\mu^\star$ to be a non-randomized policy, by noting the fact that

$$g(x,\mu) + \sum_{y=1}^{S} p_{xy}(\mu)\, h(y) = E_{u\sim\mu\left(u\middle|x\right)}\left[g(x,u) + \sum_{y=1}^{S} \rho_{xy}(u)\, h(y)\right]$$

Thus we proved Theorem 5.

Theorem 5 shows that for the purpose of finding an optimal policy, it is sufficient to restrict our attention to non-randomized stationary policy. It works as a theoretical foundation for computationally finding an optimal stationary policy, which we shall see in Section 2.5.

In practice, however, with a reasonable size of the economy this dynamic programming problem is difficult to be solved exactly given the large number of possible states and possible demand realizations; perhaps nor is it realistic or desirable to remember or store an optimal policy even if such a policy can be founded. Therefore one might prefer simpler policies that are close to optimal over an exact optimal policy. This motivates the search and study of heuristic policies in Section 2.4.

2.3 Payoff Upper Bound

Before discussing heuristic policies for the ride-sharing platform to solve the average payoff optimization problem in Section 2.4. In this Section, I provide an upper bound on the average payoff that any dynamic policy can achieve, which serves as a tool to assess the performance of heuristic policies, giving a lower bound on the relative performance in a heuristic policy compared to the optimum. The upper bound is found by solving a deterministic model.

2.3.1 A Deterministic Problem

Consider the following deterministic version of the problem: similar to our setup in Section 2.2, the traffic network has L locations and the travel time between location i and location j is t_{ij}. Time is discrete. The fleet has N vehicles, which is a *continuous* measure. In each

period t, the arrival of demand for i-to-j ride is a deterministic quantity $\lambda_{ij}(p_{ij}^t)$, which could be a fractional number.

Facing such a deterministic environment, one could still formulate the optimization problem of the ride-sharing platform as a dynamic programming problem like before, but this does not seem to simplify the problem much since a continuum of vehicles leads to an infinite system state space. However, a useful simplification is to restrict our attention to *stationary controls* and *steady states* of the system, the exact meaning of which to be specified below.

In particular, by stationary controls, we mean the following. The firm is restricted to choose a fixed set of prices forever (regardless of the system state), so $p_{ij}^t = p_{ij}$ for all $t = 0, 1, \ldots$. For a given pricing decision (p_{ij}), we know that deterministically in each period the demand for i-to-j ride is $\lambda_{ij}(p_{ij})$. The (state-independent) matching decisions of the ride-sharing platform are described by (b_{ij}) where b_{ij} is the measure of vehicles that get matched with i-to-j requests in each period, while its relocation decisions are described by (e_{ij}) where e_{ij} is the measure of vehicles that get relocated from i to j in each period. It is also required that the matching decisions and relocation decisions be fixed over time.

For the control $((p_{ij}), (b_{ij}), (e_{ij}))$ to be feasible, one immediate restriction is that

$$b_{ij} \leq \lambda_{ij}(p_{ij}) \ \ \forall i, j \tag{2.3.1}$$

i.e. the number of matches cannot exceed the demand. Secondly, for the stationary control sequence to be feasible period after period, it must be that the system is in a steady state, i.e. in each period, the number of vehicles leaving location i must be equal to the number of vehicles arriving at i:

$$\sum_j (b_{ij} + e_{ij}) = \sum_k (b_{ki} + e_{ki}), \ \forall i \tag{2.3.2}$$

Finally, we have the restriction that the total number of vehicles running in the system is N:

$$\sum_{i,j} (b_{ij} + e_{ij}) t_{ij} = N \tag{2.3.3}$$

When a control $((p_{ij}), (b_{ij}), (e_{ij}))$ satisfies the above restrictions, it corresponds to a steady system state $(N(i,\tau))_{(i,\tau)\in C}$ (where $N(i,\tau)$ is the measure of vehicles in state (i,τ)) such

that starting from this state, the stationary control is feasible and would let the system stay at this state forever. In fact, we have

$$N(i,\tau) = \sum_{k:t_{ki}-1\geq\tau} (b_{ki} + e_{ki}), \ \forall (i,\tau) \in C$$

because every vehicle in state (i,τ) must come from some location k.

Under the stationary control $((p_{ij}), (b_{ij}), (e_{ij}))$ and starting with the corresponding steady state, the per-period payoff for the ride-sharing platform is $\sum_{i,j} p_{ij} b_{ij}$, which is also the average payoff for the ride-sharing platform. Assuming that the ride sharing platform is restricted to choose among stationary controls and is allowed to start with the corresponding steady state, the ride sharing platform's problem becomes:

$$J^D \stackrel{def}{=} \max_{((p_{ij}),(b_{ij}),(e_{ij}))} \sum_{i,j\in\{1,\ldots,L\}} p_{ij} b_{ij}$$

s.t.

$$b_{ij} \leq \lambda_{ij}(p_{ij}) \ \ \forall i,j$$

$$\sum_j (b_{ij} + e_{ij}) = \sum_k (b_{ki} + e_{ki}) \ \ \forall i$$

$$\sum_{i,j} (b_{ij} + e_{ij}) t_{ij} = N$$

$$p_{ij} \geq 0, b_{ij} \geq 0, e_{ij} \geq 0 \ \forall i,j$$

An optimal solution of the deterministic problem exists because this is a convex optimization problem. Numerical solution can be obtained quite efficiently and reliably. Note that optimal prices and busy rates are unique, but optimal empty rates may not be unique when N is large.

The sub-section below shows that the average payoff that the ride-sharing platform achieves in the deterministic problem provides an upper bound on the average payoff that can be achieved by any dynamic policy in the original stochastic problem.

2.3.2 The Deterministic Payoff as an Upper Bound

Intuitively, one would expect that the uncertainty in demand in the original stochastic problem results in lower expected payoff because in the stochastic problem supply and demand are equal only in expectation but not exactly. The following theorem formalizes this idea.

Theorem 6. *If the demand in the traffic network is regular, then*

$$J^{\star}(x_0) \leq J^D, \ \forall x_0 \in X.$$

Proof. By Theorem 5, we only need to show that for any admissible non-randomized stationary policy μ, we have

$$J_{\mu}(x_0) \leq J^D, \ \forall x_0 \in X.$$

Now fix any admissible non-randomized stationary policy μ in the original problem. Recall from (2.2.4) that

$$J_{\mu}(x_0) = \rho^T g_{\mu} = \sum_{x \in \{1,\ldots,S\}} \rho_x g(x, \mu(x))$$

where ρ_x is the long-term frequency of state x under μ. Since μ is non-randomized, for any state x, $\mu(x)$ is simply some control (instead of a lottery of controls), so with some abuse of notation we can write $\mu(x) = ((p_{ij}(x)), (b_{ij}(x,\omega)), (e_{ij}(x,\omega))) \in U(x)$, thus they payoff

$$g(x, \mu(x)) = E\left[\sum_i \sum_j p_{ij}(x) b_{ij}(x,\omega) \middle| x\right] = \sum_i \sum_j p_{ij}(x) b_{ij}(x)$$

where $b_{ij}(x) \stackrel{def}{=} E\left[b_{ij}(x,\omega) \middle| x\right]$ is the expected number of busy vehicles on direction i-to-j given state x and stationary policy μ. I suppress the dependence on μ in the notation $b_{ij}(x)$ because we have fixed μ at the beginning of our discussion.

Similarly let $e_{ij}(x) \stackrel{def}{=} E\left[e_{ij}(x,\omega) \middle| x\right]$. Thus $e_{ij}(x)$ is the expected number of vehicles that are relocated from i to j when the current state is x and we adopt policy μ. Again I suppress the dependence on μ.

It can be shown that the following conditions hold (see proof in Appendix):

$$b_{ij}(x) \leq \lambda_{ij}(p_{ij}(x)), \ \forall x \in X = \{1, ..., S\}, \ \forall i, j \in \{1, ..., L\} \tag{2.3.4}$$

$$\sum_x \rho_x \sum_j (b_{ij}(x) + e_{ij}(x)) = \sum_x \rho_x \sum_k (b_{ki}(x) + e_{ki}(x)), \ \forall i \in \{1, ..., L\} \tag{2.3.5}$$

$$\sum_x \rho_x \left(\sum_i \sum_j (b_{ij}(x) + e_{ij}(x)) t_{ij} \right) = N \tag{2.3.6}$$

The conditions can be understood intuitively. (2.3.4) is straightforward: it says that under any state, the expected number of i-to-j matches cannot exceed the expected demand, which is trivially true given that μ is an admissible policy. (2.3.5) says that the long-run average expected number of vehicles flowing out of location i is equal to the long-run average number of vehicles flowing into location i, which is intuitively true. (2.3.6) says roughly that on average the number of vehicles sent times the traveling time gives the total number of vehicles.

Now consider the following "relaxed" deterministic problem:

$$J^{RD} \stackrel{def}{=} \max_{((p_{ij}(x)),(b_{ij}(x)),(e_{ij}(x)))} \sum_{x \in \{1,...,S\}} \rho_x \left(\sum_{i,j \in \{1,...,L\}} p_{ij}(x) b_{ij}(x) \right)$$

s.t.

$$b_{ij}(x) \leq \lambda_{ij}(p_{ij}(x)), \ \forall x, i, j$$

$$\sum_x \rho_x \sum_j (b_{ij}(x) + e_{ij}(x)) = \sum_x \rho_x \sum_k (b_{ki}(x) + e_{ki}(x)), \ \forall i$$

$$\sum_x \rho_x \left(\sum_i \sum_j (b_{ij}(x) + e_{ij}(x)) t_{ij} \right) = N$$

$$p_{ij}(x) \geq 0, b_{ij}(x) \geq 0, e_{ij}(x) \geq 0, \ \forall x, i, j$$

Then clearly, $J_\mu(x_0) \leq J^{RD}$. Now to show that $J_\mu(x_0) \leq J^D$, we only need to show that $J^{RD} \leq J^D$.

In fact the opposite direction is obvious: $J^{RD} \geq J^D$, because if $\{p_{ij}, b_{ij}, e_{ij}\}$ attains the maximum of the deterministic problem, then simply set $p_{ij}(x) = p_{ij}$, $b_{ij}(x) = b_{ij}$,

$e_{ij}(x) = e_{ij}$ for all x, note that all the constraints in the relaxed problem are satisfied, and we can immediately see that $J^{RD} \geq J^{D}$. But it turns out that the relaxed problem does not yield strictly higher value because of the concavity of the revenue function. To prove this, first note that the first set of constraints $b_{ij}(x) \leq \lambda_{ij}(p_{ij}(x))$ is always binding. So we may as well write:

$$J^{RD} = \max_{((b_{ij}(x)),(e_{ij}(x)))} \sum_{x \in \{1,\ldots,S\}} \rho_x \left(\sum_{i \neq j \in \{1,\ldots,L\}} r_{ij}(b_{ij}(x)) \right)$$

s.t.

$$\sum_x \rho_x \sum_j (b_{ij}(x) + e_{ij}(x)) = \sum_x \rho_x \sum_k (b_{ki}(x) + e_{ki}(x)), \ \forall i$$

$$\sum_x \rho_x \left(\sum_i \sum_j (b_{ij}(x) + e_{ij}(x)) t_{ij} \right) = N$$

$$b_{ii}(x) = 0, e_{ii}(x) \geq 0 \ \forall i$$

$$b_{ij}(x) \geq 0, e_{ij}(x) \geq 0 \ \forall i \neq j$$

where $r_{ij}(\cdot)$ is the revenue rate function defined before. Similarly for the deterministic problem we may rewrite it as:

$$J^{D} = \max_{((b_{ij}),(e_{ij}))} \sum_{i \neq j \in \{1,\ldots,L\}} r_{ij}(b_{ij}(x))$$

s.t.

$$\sum_j (b_{ij} + e_{ij}) = \sum_k (b_{ki} + e_{ki}) \ \forall i$$

$$\sum_{i,j} (b_{ij} + e_{ij}) t_{ij} = N$$

$$b_{ii} = 0, e_{ii} \geq 0 \ \forall i$$

$$b_{ij} \geq 0, e_{ij} \geq 0 \ \forall i \neq j$$

Suppose $\left(\left(b^{\star}_{ij}(x)\right), \left(e^{\star}_{ij}(x)\right)\right)$ attains the maximum of the relaxed deterministic problem (the existence of a maximizer is guaranteed by the continuity of the objective function and the fact that the choice set is compact). Then let $b'_{ij}(x) = b'_{ij} \stackrel{def}{=} \sum_x \rho_x b^{\star}_{ij}(x)$, $e'_{ij}(x) =$

$e'_{ij} \stackrel{def}{=} \sum_x \rho_x e^\star_{ij}(x)$ for all x. Obviously $\left(\left(b'_{ij}(x)\right), \left(e'_{ij}(x)\right)\right)$ also satisfies the constraints of the relaxed problem; more importantly, because $r_{ij}(\cdot)$ is concave, we immediately see that

$$\sum_x \rho_x \left(\sum_{i\neq j} r_{ij}\left(b'_{ij}(x)\right)\right) = \sum_{i\neq j} r_{ij}\left(b'_{ij}\right) = \sum_{i\neq j} r_{ij}\left(\sum_x \rho_x b^\star_{ij}(x)\right) \geq \sum_{i\neq j}\sum_x \rho_x r_{ij}\left(b^\star_{ij}(x)\right) = J^{RD}$$

Therefore we must have that $\sum_x \rho_x \left(\sum_{i\neq j} r_{ij}\left(b'_{ij}(x)\right)\right) = J^{RD}$. Note also that $\left(\left(b'_{ij}\right), \left(e'_{ij}\right)\right)$ satisfies constraints of the deterministic problem, so $\sum_{i\neq j} r_{ij}\left(b'_{ij}\right) \leq J^D$. Thus we prove that $J^{RD} \leq J^D$. □

Theorem 6 says that the optimal average payoff in a deterministic model provides an upper bound on the average payoff that can be obtained by any dynamic policy in the stochastic model. This upper bound can be used to construct a lower bound on the relative performance of heuristic policies discussed in the following section relative the optimum.

Notice that in the proof of Theorem 6, the crucial assumption is the concavity of the revenue function, and thus Poisson distribution assumption of the ride requests is not necessary for the result to hold, and thus the results can be general to distributions for ride requests as long as they induce concave revenue functions.

2.4 Heuristic Policies

We next examine heuristic policies based on solutions to the deterministic problem. The idea of those heuristics is to "mimic" the system behavior of the deterministic solution, with the intuition that if the market is large or thick enough, then in a normalized sense, the stochasticity in the system becomes almost deterministic, thus we would expect the heuristics to have approximately the same performance as the optimal deterministic solution.

Two kinds of implementations of the deterministic solution are considered: Stay-at-Steady-State (SSS) policies, and Flexible Accommodation (FA) policies. Both kinds of

heuristics fix the prices at the levels suggested by a deterministic solution for all possible system states, but they differ in their allocation decisions (i.e. matching and relocation). For the family of SSS policies, I provide a lower bound on its performance relative to the optimum that apply in any system within our theoretical framework. As corollaries, I show that SSS policies are asymptotically optimal in thick markets. For the family of FA policies, I conjecture that some appropriately implemented versions of them are also asymptotically optimal, and I provide simulation evidence to support this conjecture.

2.4.1 Stay-at-Steady-State Policies

The key idea of the SSS policies is for the ride-sharing platform to choose policies such that the traffic network stays at states that are "close" to the steady state corresponding to a given deterministic solution. This family of policies can be proven to be asymptotically optimal as the scale of the economy increases.

I start with a special knife-edge case in which it is possible to stay at the exact deterministic steady state in each period, in which the heuristic policy takes a simple and straightforward form. I then present some generalized versions of the heuristic that are applicable to more general demand environment. Even though the generalized versions seem complicated at the first glance, the complications come merely from the fact that the stochastic system is discrete in the number of vehicles, which essentially brings the complexity of integer programming.

Special Cases: Strictly-Stay-at-Steady-State Policies

A deterministic solution $\left(\left(p_{ij}^{\star}\right), \left(b_{ij}^{\star}\right), \left(e_{ij}^{\star}\right)\right)$ corresponds to a steady state $x^{\star} = \left(N^{\star}(i,\tau)\right)_{(i,\tau)\in C}$:

$$N^{\star}(i,\tau) = \sum_{k:t_{ki}-1\geq\tau} \left(b_{ki}^{\star} + e_{ki}^{\star}\right), \ \forall (i,\tau) \in C$$

Generically, this steady state involves non-integer amount of vehicles, which is not feasible in our traffic network as the number of vehicles is discrete. However, if $b_{ki}^{\star} + e_{ki}^{\star}$ happens to be an integer for all i and k, then this steady state is indeed a state in the discrete system. Furthermore, the vehicle flows in the deterministic solution can be replicated exactly. This

inspires us to consider the following heuristic:

A Strictly-Stay-at-Steady-State (S-SSS) Policy *Suppose* $\left(\left(p_{ij}^{\star}\right), \left(b_{ij}^{\star}\right), \left(e_{ij}^{\star}\right)\right)$ *is an optimal solution in the deterministic problem, and assume that* $b_{ki}^{\star} + e_{ki}^{\star}$ *is an integer for all* i *and* k*. Let* $x^{\star} = \left(N^{\star}(i,\tau)\right)_{(i,\tau)\in C}$ *be the corresponding steady state. The S-SSS heuristic corresponding to the deterministic solution* $\left(\left(p_{ij}^{\star}\right), \left(b_{ij}^{\star}\right), \left(e_{ij}^{\star}\right)\right)$ *is as follows: If the current state is not* $x^{\star}$*, relocate vehicles so that the system arrives at* $x^{\star}$ *as soon as possible*[13]*. If the current state is* $x^{\star}$*, first set prices as* $p = \left(p_{ij}^{\star}\right)$*; then for each origin-destination pair* (i,j)*, if the realized demand for* i*-to-*j *ride is* ω_{ij}*, accept* $\min\left\{\omega_{ij}, b_{ij}^{\star} + e_{ij}^{\star}\right\}$ *of them, and relocate the remaining vehicles from location* i *to* j *without riders.*

To illustrate the S-SSS policy defined above, consider the following numeric example: The network has 3 locations: $\{A, B, C\}$, and 6 cars. For simplicity assume that it takes one period to travel between any two locations. Suppose $\left(\left(p_{ij}^{\star}\right), \left(b_{ij}^{\star}\right), \left(e_{ij}^{\star}\right)\right)$ is a solution to the deterministic problem, and we illustrate the numeric values of $\left(\left(b_{ij}^{\star}\right), \left(e_{ij}^{\star}\right)\right)$ in the left panel of Figure 2-1, with variables being 0 omitted (e.g. $e_{AA}^{\star} = 0$, $b_{BB}^{\star} = 0$, $e_{AB}^{\star} = 0$) for simplicity. The exact numeric values of the prices $\left(p_{ij}^{\star}\right)$ are not crucial for our analysis and therefore are also omitted here.

It can be easily verified that indeed $b_{ij}^{\star} + e_{ij}^{\star} = 1$ (integer) in every direction, meaning that for each location, the inflow and outflow of cars are equal in each period. The steady state of the system corresponding to the deterministic solution is the state in which there are 2 cars in each location. S-SSS policy then dictates that once we are in this steady state, then in each period exactly one car, carrying a rider if there is a request or staying empty otherwise, travels from one location to another. In particular, if the realization of demands (ω_{ij}) is as shown on the right panel of Figure 2-2, and if we begin with the steady state, then one of the two requests from C to B will be rejected according to S-SSS policy, even though location C has two available cars to begin with and there is no request on the C-to-A direction. S-SSS policy requires that we send an empty car from C to A instead of being

[13]There is a bit of ambiguity for the exact implementation of the policy when starting from a non-steady state. However, note that it takes at most $2max_{i,j}t_{ij}$ periods to deterministically arrive at $x^{\star}$, and since the pricing and matching decisions in a finite number of periods do not matter for the system's average payoff over the infinite horizon, the pricing and matching decisions in S-SSS policy for the stage before arriving at steady state are not important, and therefore can be arbitrary.

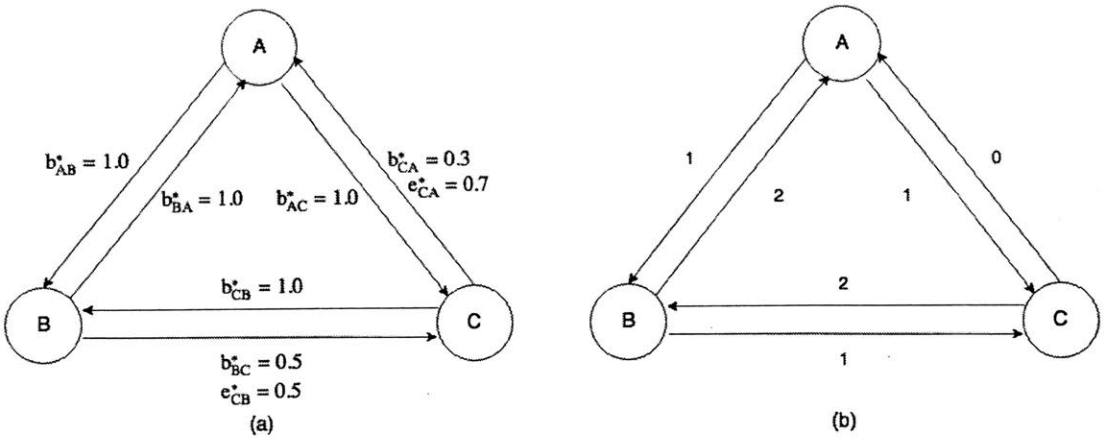


Figure 2-1. A 3-location, 6-car network, with unit travel time on all directions
The left panel (a) illustrates a deterministic solution (matching and relocation rates being 0 and prices omitted). The right panel (b) shows a possible realization of demands.

flexible and let that car serve a C-to-B request instead.

Therefore, S-SSS policy keeps the system exactly at the "desired" state - the steady state corresponding to the deterministic optimal solution, even if doing so takes away some dispatching flexibility by requiring seemingly unnecessary rejections occasionally.

By sticking to the steady state, the expected average payoff of S-SSS heuristic is very straightforward. Recall from (2.2.4) that the formula for the expected average payoff of a stationary policy μ is simply the weighted average payoff per period with the weights being the long-run stationary distribution:

$$J_\mu(x_0) = \rho^T g_\mu = \sum_{x \in \{1,\ldots,S\}} \rho_x g(x, \mu(x))$$

Note that S-SSS policy is a stationary policy under which the long-run probability of $x^\star$ is 1 while all other system states have zero long-run probability. Therefore, let $J^{S-SSS}(x_0)$ denote the expected average payoff under the corresponding S-SSS policy starting at state

x_0, then we have

$$\begin{aligned} J^{S-SSS}(x_0) &= g(x^\star, \mu_{S-SSS}(x^\star)) \\ &= \sum_{i,j} p_{ij}^\star E_\omega \left[b_{ij}(\omega) \middle| x^\star, \mu_{S-SSS} \right] \\ &= \sum_{i,j} p_{ij}^\star E_\omega \left[\omega_{ij} - \left(\omega_{ij} - \left(b_{ij}^\star + e_{ij}^\star \right) \right)^+ \middle| x^\star, \mu_{S-SSS} \right] \end{aligned}$$

where $x^+ \stackrel{def}{=} \max(x, 0)$ represents the positive part of x.

The following theorem provides a lower bound on the performance of S-SSS relative to the optimal policy

Theorem 7. *Let $\left(\left(p_{ij}^\star \right), \left(b_{ij}^\star \right), \left(e_{ij}^\star \right) \right)$ be a deterministic optimal solution and let $x^\star$ be its corresponding steady state. Assume that $b_{ij}^\star$ and $e_{ij}^\star$ are all integers. Let $J^{S-SSS}(x_0)$ denote the expected average payoff under the corresponding S-SSS policy starting at state x_0. Then*

$$\frac{J^{S-SSS}(x_0)}{J^\star(x_0)} \geq \frac{J^{S-SSS}(x_0)}{J^D} \geq 1 - \frac{1}{2} \left(\frac{\sum_{i,j} p_{ij}^\star \sqrt{b_{ij}^\star}}{\sum_{i,j} p_{ij}^\star b_{ij}^\star} \right).$$

Proof. From the analysis above we have the expected average payoff for the S-SSS policy being

$$J^{S-SSS}(x_0) = \sum_{i,j} p_{ij}^\star E_\omega \left[\omega_{ij} - \left(\omega_{ij} - \left(b_{ij}^\star + e_{ij}^\star \right) \right)^+ \middle| x^\star, \mu_{S-SSS} \right].$$

Gallego (1992) shows that for any random variable D with finite mean μ and finite standard deviation σ, and for any real number d,

$$E\left[(D-d)^+ \right] \leq \frac{\sqrt{\sigma^2 + (d-\mu)^2} - (d-\mu)}{2}$$

Now applying this inequality and noting that $\omega_{ij} \middle| x^\star, \mu_{S-SSS} \sim Poisson\left(b_{ij}^\star \right)$, we obtain

$$E_\omega \left[\left(\omega_{ij} - \left(b_{ij}^\star + e_{ij}^\star \right) \right)^+ \middle| x^\star, \mu_{S-SSS} \right] \leq \frac{\sqrt{b_{ij}^\star + (e_{ij}^\star)^2} - e_{ij}^\star}{2} \leq \frac{\sqrt{b_{ij}^\star}}{2}$$

Thus

$$\begin{aligned} J^{S-SSS}(x_0) &= \sum_{i,j} p_{ij}^\star E_\omega \left[\omega_{ij} - \left(\omega_{ij} - \left(b_{ij}^\star + e_{ij}^\star\right)\right)^+\right] \\ &\geq \sum_{i,j} p_{ij}^\star b_{ij}^\star - \frac{1}{2}\sum_{i,j} p_{ij}^\star \sqrt{b_{ij}^\star} \end{aligned}$$

Noting that $J^D = \sum_{i,j} p_{ij}^\star b_{ij}^\star$, we immediately prove the second inequality in the theorem. The first inequality simply follows from Theorem 6. □

To understand the significance of Theorem 7 more easily, consider a sequence of problems indexed by integers $K = 1, 2, ...$ in which the scale of economy increases linearly, meaning the total supply of vehicles $N_K = KN$ and demand functions $\lambda_{ij-K}(p_{ij}) = K\lambda_{ij}(p_{ij})$ for some fixed N and $\lambda_{ij}(\cdot)$, while the number of locations L and the travel times t_{ij}'s remain the same for all problems in the sequence. This generates a sequence of problems with proportionately larger supply and demand. It is not hard to see that this sequence of problems has the same optimal deterministic prices $\left(p_{ij}^\star\right)$ for all K, and if $\left(\left(p_{ij}^\star\right), \left(b_{ij}^\star\right), \left(e_{ij}^\star\right)\right)$ is an optimal solution to the deterministic problem for $K = 1$ and $\left(\left(p_{ij}^\star\right), \left(Kb_{ij}^\star\right), \left(Ke_{ij}^\star\right)\right)$ is an optimal solution to the deterministic problem for problem K. Let $J_K^{S-SSS}(x_0)$ denote the expected average payoff of S-SSS policy corresponding to $\left(\left(p_{ij}^\star\right), \left(Kb_{ij}^\star\right), \left(Ke_{ij}^\star\right)\right)$ in problem K, we have

$$\begin{aligned} \frac{J_K^{S-SSS}(x_0)}{J_K^\star(x_0)} &\geq 1 - \frac{1}{2}\left(\frac{\sum_{i,j} p_{ij}^\star \sqrt{Kb_{ij}^\star}}{\sum_{i,j} p_{ij}^\star Kb_{ij}^\star}\right) \\ &= 1 - O\left(K^{-1/2}\right) \end{aligned}$$

This shows that the SSS heuristic is asymptotically optimal as the scale of the problem increases ($K \to \infty$).

The theorem can be generalized if we want to relax the Poisson distribution assumption of demands. More generally, if under price $p_{ij}^\star$ the demand ω_{ij} is a random variable with

mean $b^\star_{ij}$ and standard deviation σ_{ij} (for all i, j), we have

$$\begin{aligned}
\frac{J^{S-SSS}(x_0)}{J^D} &\geq 1 - \frac{1}{2}\left(\frac{\sum_{i,j} p^\star_{ij}\sigma_{ij}}{\sum_{i,j} p^\star_{ij} b^\star_{ij}}\right) \\
&= 1 - \frac{1}{2}\left(\frac{\sum_{i,j}\left[p^\star_{ij} b^\star_{ij} \cdot \frac{\sigma_{ij}}{b^\star_{ij}}\right]}{\sum_{i,j} p^\star_{ij} b^\star_{ij}}\right) \\
&= 1 - \frac{1}{2}\sum_{i,j}\left[\frac{p^\star_{ij} b^\star_{ij}}{\sum_{i,j} p^\star_{ij} b^\star_{ij}} \cdot \frac{\sigma_{ij}}{b^\star_{ij}}\right]
\end{aligned}$$

Therefore if either $\frac{\sigma_{ij}}{b^\star_{ij}}$ is negligibly small in all directions (which is the case for the $K \to \infty$ economy as discussed above), or those directions whose $\frac{\sigma_{ij}}{b^\star_{ij}}$ is not negligibly small have no major impact (e.g. their $\frac{\sigma_{ij}}{b^\star_{ij}}$ is bounded and their contribution to total sale $\frac{p^\star_{ij} b^\star_{ij}}{\sum_{i,j} p^\star_{ij} b^\star_{ij}}$ are all negligibly small), then we know that S-SSS heuristic would have performance close to its deterministic counterpart. The interpretation of the result is straight-forward, when the stochasticity brought by the random demand is small compared to the scale of matching rides in the deterministic problem, the sub-optimality introduced by the randomly unaccommodated rides is small.

General Cases: Generalized-Stay-at-Steady-State Policies

There are two shortcomings of the aforementioned S-SSS policy: firstly (and most importantly), it does not exist in general because it requires that $b^\star_{ij} + e^\star_{ij}$ to be an integer for all i and j; secondly, even though the control to take when the system is already in the steady state is well-defined, it is a bit ambiguous about the exact control to take when the system is not in the steady state yet (though this problem is arguably not crucial and is probably fixable by more careful design).

The above shortcomings can be fixed quite easily with some slight modifications. When $b^\star_{ij} + e^\star_{ij}$ is not an integer, it is impossible to send fractional amount of cars on that direction, but we can try to send, for example, $\left\lfloor b^\star_{ij} + e^\star_{ij} \right\rfloor$ cars to location j in total (where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x), and accommodate as many requests from i to j given this "total target flow". But we need to deal with the following two questions: What if there are more than enough cars available at location i such that, after sending $\left\lfloor b^\star_{ij} + e^\star_{ij} \right\rfloor$

cars from i to j for all j, there are still cars left? And what if there are not enough cars available at location i such that it is infeasible to send $\left\lfloor b_{ij}^{\star} + e_{ij}^{\star} \right\rfloor$ cars to all location j?

The ways to deal with those questions are not unique; the proposed policy below is just one possible solution, but we study it extensively in order to be concrete.

A Generalized-Stay-at-Steady-State (G-SSS) Policy. *Suppose that* $\left(\left(p_{ij}^{\star}\right), \left(b_{ij}^{\star}\right), \left(e_{ij}^{\star}\right)\right)$ *is a deterministic optimal solution. Define* $d_{ij}^{\star} = \left\lfloor b_{ij}^{\star} + e_{ij}^{\star} \right\rfloor$ *(here* $\lfloor x \rfloor$ *denotes the largest integer less than or equal to* x*) to be the "i-to-j target". The G-SSS policy corresponding to the above deterministic optimal solution is a stationary policy that works in the following way: Set prices at* $p^{\star}$ *for all states. Suppose the current state is* $x = (N(i,\tau))_{(i,\tau)\in C}$*. First we define the "dispatching quota"* $q_{ij}(x)$ *for all i-to-j directions as follows: If* $N(i,0) \geq \sum_j d_{ij}^{\star}$*, then* $q_{ij}(x) = d_{ij}^{\star}$*; otherwise, first set* $q_{ij}(x) = d_{ij}^{\star}$*, then starting from* $j = 1$ *to* $j = L$*, sequentially lower* $q_{ij}(x)$ *by 1 (skip this step if* $q_{ij}(x)$ *is already zero), repeat for more rounds if needed, until when* $\sum_j q_{ij}(x) = N(i,0)$*. Then, the contingent dispatching plan* $(b_{ij}(x,\omega), e_{ij}(x,\omega))$ *is as follows:*

$$b_{ij}(x,\omega) = \min\{\omega_{ij}, q_{ij}(x)\} \quad \forall i,j$$

$$e_{ij}(x,\omega) = q_{ij}(x) - b_{ij}(\omega) \quad \forall i, \forall j \neq i$$

$$e_{ii}(x,\omega) = N(i,0) - \sum_{j\neq i}\left(b_{ij}(\omega) + e_{ij}(\omega)\right) \quad \forall i$$

that is, for $j \neq i$*, always dispatch* $q_{ij}(x)$ *cars from* i *to* j*, while accommodate as many i-to-j requests given this constraint; keep the remaining* $(i,0)$ *cars at the same place (i.e. at location i), if there is still any.*

To illustrate how the above G-SSS policy works, consider the following numeric example.

Example. Consider a 3-location ($\{A, B, C\}$) network with 12 cars, and for simplicity assume that it takes one period to travel from any origin to any destination. Suppose the demands of the city are such that $\left(\left(p_{ij}^{\star}\right), \left(b_{ij}^{\star}\right), \left(e_{ij}^{\star}\right)\right)$ is a solution to the deterministic problem; we illustrate the numeric values of $\left(\left(b_{ij}^{\star}\right), \left(e_{ij}^{\star}\right)\right)$ in the left panel of Figure 2-2, with variables being 0 omitted (e.g. $e_{AA}^{\star} = 0$, $e_{AC}^{\star} = 0$, $b_{AA}^{\star} = 0$). The exact numeric values of the prices

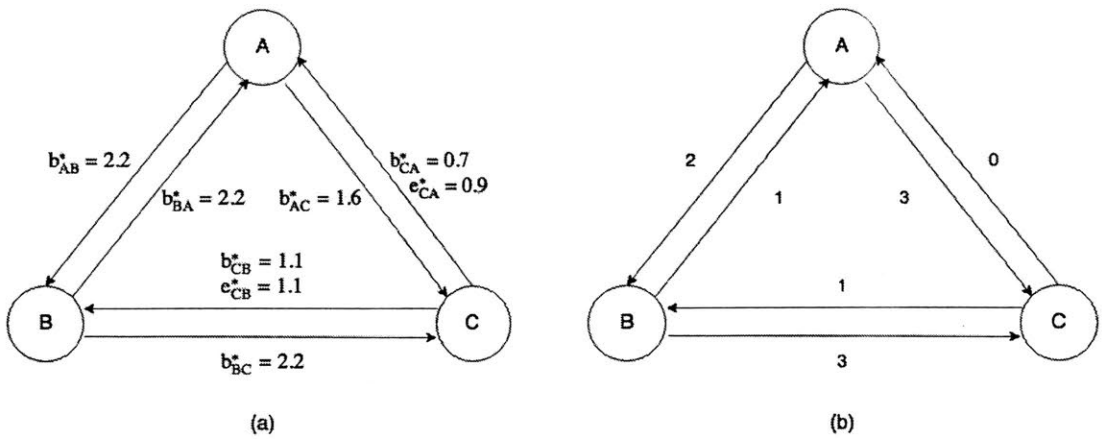


Figure 2-2. A 3-location, 12-car network, with unit travel time on all directions
The left panel (a) illustrates a deterministic solution (matching and relocation rates being 0 and prices omitted). The right panel (b) shows a possible realization of demands.

$\left(p^\star_{ij}\right)$ are not crucial for our analysis and therefore are also omitted here.

Under G-SSS policy, we set prices at $\left(p^\star_{ij}\right)$ regardless of the system state. The target flow from A to B is $d^\star_{AB} = \lfloor b^\star_{AB} + e^\star_{AB} \rfloor = \lfloor 2.2 + 0 \rfloor = 2$; similarly, we have $d^\star_{AC} = 1$, $d^\star_{AA} = 0$, $d^\star_{BA} = 2$, etc. Now suppose we are in the state in which there are 2, 4, 6 cars in location A, B, C respectively (and we denote this state as $x = (2, 4, 6)$). According to the policy, we need to have some (arbitrary but fixed) rank order of the locations, so suppose we number location A, B, C as location 1, 2, 3 respectively. Then, for location A, since $d^\star_{AA} + d^\star_{AB} + d^\star_{AC} = 0 + 2 + 1 = 3 > 2$, the dispatching quota needs to be adjusted by lowering the targets by one at a time sequentially; in this case, we find that eventually, $q_{AA}(x) = 0$, $q_{AB}(x) = 1$, $q_{AC}(x) = 1$ works. For location B and C, we see that there are enough cars in each location, so the quotas are just equal to the targets: $q_{BA}(x) = d^\star_{BA} = 2$, $q_{BB}(x) = d^\star_{BB} = 0$, $q_{BC}(x) = d^\star_{BC} = 2$; $q_{CA}(x) = d^\star_{CA} = 1$, $q_{CB}(x) = d^\star_{CB} = 2$, $q_{CC}(x) = d^\star_{CC} = 0$.

Once the quotas are figured out, it is straightforward to see what should be the dispatching plan given any realized demand. For example, suppose the realized demand is as shown in the right panel of Figure 2-2. Then only 1 of the 2 requests from A to B will be accepted, 1 of the 3 requests from A to C will be accepted, 2 of the 3 requests from B to C will be accepted, and the rest of the requests in other directions will all be accepted. 1 car is relocated from B to A, 1 car is relocated from C to A, 1 car is relocated from C to B, and 3 cars stays in C.

Note that regardless of the demand realization, under G-SSS policy, the system state in the next period must be $(3,3,6)$. In fact, one can check that regardless of the demand realizations in the future, the system state should be alternating between $(2,4,6)$ and $(3,3,6)$ thereafter.

What if we start with some different initial states? We can do the following thought experiments; note that the system state evolution under G-SSS policy does not depend on demand realization and is deterministic:

$(6,3,3) \to (5,4,3) \to (5,4,3) \to ...$

$(4,4,4) \to (4,4,4) \to (4,4,4) \to ...$

$(12,0,0) \to (9,2,1) \to (7,3,2) \to (5,4,3) \to (5,4,3) \to ...$

From the above "thought experiments" we can make the following rough observations about G-SSS policy: firstly, the long-run frequency of states may depend on the initial state; secondly, the states that would appear in the long run cannot be "too extreme". For example, intuitively we should expect that the state $(12,0,0)$ should not return in the long run.

In fact, we can formalize the second observation by the following lemma (with the addition of some mild assumptions), which basically says that the (discrete) states that can possibly have positive long-run probability under G-SSS policy are "not far from" the (continuous) steady state. For ease of notation, in the following results we assume $t_{ij} = 1$ for all i and j, but all the results can be generalized to cases with general travel times.

Lemma 14. *Assume that $t_{ij} = 1$ for all $i, j \in \{1, ..., L\}$. $\left(\left(p_{ij}^{\star}\right), \left(b_{ij}^{\star}\right), \left(e_{ij}^{\star}\right)\right)$ is a deterministic optimal solution, and assuming that $b_{ij}^{\star} + e_{ij}^{\star} > L$ for all i and $j \neq i$,*[14]*, if system state $x = (N(i,0))_{(i,0)\in C}$ has positive long-run probability under the G-SSS policy, it must satisfy that for all i,*

$$\sum_{k} \left(b_{ki}^{\star} + e_{ki}^{\star}\right) - \left(3L^2 - 1\right) \leq N(i,0) < \sum_{j} \left(b_{ij}^{\star} + e_{ij}^{\star}\right) + (3L+1)(L-1).$$

The details of the proof of the above lemma is deferred to Appendix.

[14]This condition is not necessary to obtain the qualitative result that G-SSS is asymptotically optimal; for that purpose it can be relaxed to the weaker condition of, say, $\Sigma_j (b_{ij}^{\star} + e_{ij}^{\star}) > L$, with slight modifications of the bounds correspondingly. But assuming it simplifies the proof.

Based on the above result, we can soon generalize the asymptotic optimality result to G-SSS policy, as follows (see Appendix for proof details).

Theorem 8. *Assume that $t_{ij} \equiv 1$. Let $\left(\left(p_{ij}^{\star}\right), \left(b_{ij}^{\star}\right), \left(e_{ij}^{\star}\right)\right)$ be a deterministic optimal solution, and assume that $b_{ij}^{\star} + e_{ij}^{\star} > L$ for all i and $j \neq i$,. Let $J^{G-SSS}(x_0)$ denote the expected average payoff under the corresponding G-SSS policy starting at state x_0. Then*

$$\frac{J^{G-SSS}(x_0)}{J^{\star}(x_0)} \geq \frac{J^{G-SSS}(x_0)}{J^{D}} \geq 1 - \frac{\sum_{i,j} p_{ij}^{\star}\left(\frac{\sqrt{b_{ij}^{\star}}}{2} + 3L + 2\right)}{\sum_{i,j} p_{ij}^{\star} b_{ij}^{\star}}.$$

Similar to our discussions about S-SSS, consider a sequence of problems indexed by integers $K = 1, 2, ...$, with total supply of vehicles being $N_K = KN$ and demand functions being $\lambda_{ij-K}(p_{ij}) = K\lambda_{ij}(p_{ij})$ for the Kth problem, where N and $\lambda_{ij}(\cdot)$ are the base number of vehicles and demand functions for the 1th first problem. The above theorem implies that

$$\frac{J_K^{G-SSS}(x_0)}{J_K^{\star}(x_0)} \geq 1 - O\left(K^{-\frac{1}{2}}\right)$$

which mean that the G-SSS heuristic policy is asymptotically optimal as $K \to \infty$.

Here we highlight some key features of the G-SSS policy:

1. As in S-SSS policy, under G-SSS the movement of cars in the system is not affected by the realization of demands, i.e. $b_{ij}(\omega) + e_{ij}(\omega)$ does not depend on the realization of ω. In other words, once we know the initial state in $t = 0$, we know all the future states deterministically. The realization of demands only affects the realization of revenue/payoff, but not the state.

2. When there are sufficient number of cars in location i, we mimic the deterministic car flow by sending $\left\lfloor b_{ij}^{\star} + e_{ij}^{\star} \right\rfloor = d_{ij}^{\star}$ cars from i to $j \neq i$ and leave the rest in i. When there are not sufficient cars, we are unable to send as many cars, but we strive to keep the "deficit" balanced in all the directions.

2.4.2 Flexible Accommodation Heuristics

As illustrated the G-SSS heuristic requires "seemingly unnecessary rejection of requests" occasionally because the ride-sharing platform only cares about achieving steady state under G-SSS. This feature is undesirable if one is concerned about even slight discounting of the future, or potential mis-specification of the model. I therefore consider another kind of heuristic which tries to accommodate as many ride requests as possible.

The main idea of this kind of policy is as follows:

1. Fix the price at the deterministic optimal price level;

2. Once demands are realized, instead of trying to keep the system at (or close to) "steady state" as in the G-SSS policy, now the priority is to accommodate as many requests as possible for each location i. If there are more than sufficient number of cars at location i to accommodate all requests from i, then we are left with unmatched cars to be relocated in the next step. If there are fewer cars available at location i than requested, simply randomly match cars with riders;

3. After step 2, if we are left with unmatched cars at location i, we relocate those cars in a way such that the total car flow out of location i "mimics" the solution in the deterministic problem. The way to "mimic" the deterministic solution is not unique. For example, one way to do it rather "loosely" is to independently assign each unmatched car at location i to some relocation destination, where destinations that have higher "deficit" flow ("deficit" flow defined as target total flow minus matched flow) receive assignments with higher probability - see the example below for the exact determination of those probabilities. I call this policy ***FA-mdR policy***, where FA stands for "flexible accommodation" and mdR stands for "mimicking deterministic relocation". [15]

Also in the spirit of flexible accommodation, I consider two more heuristics that are "careless"

[15] We could also do the "mimicking" more carefully: for example, if we have more than sufficient number of cars to "make up for the deficits", we first make sure that they are relocated so that all the deficits are made up to, and then randomly assign the rest of the cars, while if we don't have sufficient cars to meet the deficits, we just randomly match the remaining cars to the deficit spots. In the simulation, I find the convergence property of this policy is very similar to that of FA-mdR.

in terms of relocation, which I call ***FA-rR*** (flexible accommodation with random relocation) and ***FA-woR*** (flexible accommodation without relocation) respectively. Both policies accommodate demands in the same way as the previous FA policies, but when it comes to relocation, FA-rR randomly assigns any remaining cars to any direction with equal probability, while FA-woR policy just doesn't relocate at all (that is, all the remaining cars that are not assigned any rider simply stay at the same location). In other words, FA-rR and FA-woR policies do not use information in the deterministic solution other than the prices.

Before studying the performance of those heuristics, we first use a concrete example to explain how those heuristics work.

Example. Consider a 3-location ($\{A, B, C\}$) network with 12 cars, and for simplicity assume that it takes one period to travel between any two locations. Suppose the demands of the traffic network is regular and $\left(\left(p_{ij}^{\star}\right), \left(b_{ij}^{\star}\right), \left(e_{ij}^{\star}\right)\right)$ is a solution to the deterministic problem; we illustrate the numeric values of $\left(\left(b_{ij}^{\star}\right), \left(e_{ij}^{\star}\right)\right)$ in the left panel of Figure 2-2, with variables being 0 omitted (e.g. $e_{AA}^{\star} = 0$, $e_{AC}^{\star} = 0$). The exact numeric values of the prices $\left(p_{ij}^{\star}\right)$ are not crucial for our analysis and therefore are also omitted here.

For the purpose of illustration, consider the system state in which there are 4 (available) cars in each location. For a given deterministic solution, all the heuristics proposed in this section (including G-SSS policy in the previous subsection) simply propose to set the prices at $\left(p_{ij}^{\star}\right)$, the deterministic optimal prices, regardless of the system state; they differ in their contingent plan of matching and relocation. Consider a possible realization of demands as illustrated on the right panel of Figure 2-2. How would different heuristics differ in their matching and relocation decisions?

First consider the 4 cars at location A. There are 2 ride requests from A to B and 3 ride requests from A to C, so it is impossible to accommodate all those requests. All the heuristics proposed in this subsection (FA-mdR, FA-rR, FA-woR) give the same (randomized) allocation decision for those cars, that is, randomly match the cars with requests. Thus each of the 5 requests that originate from A may end up unserved with equal probability, and exactly one of them will be unserved. (In contrast, under G-SSS policy, both requests from A to B will be served, only one of the requests from A to C will be served, and one car will remain at location A even though it is available to serve one of the rejected requests from A

to C).

Next consider location B. There are 3 requests from B to C and 1 request from B to A, so under any of the heuristics in this subsection, all those requests will be served, and no cars will be left unmatched. (In contrast, under G-SSS policy, one request from B to C will be rejected, and one empty car will be relocated from B to A).

Finally consider location C. There is only one request that originates from C in this period, while there are four available cars. Naturally this request will be served under any policy (including G-SSS). However, for the remaining 3 cars, different heuristics have different relocation decisions:

- FA-mdR: each of the 3 cars will be independently relocated to location B with probability $\frac{(2.2-1)}{(2.2-1)+(1.6-0)} = 0.43$ and to location A with probability $\frac{(1.6-0)}{(2.2-1)+(1.6-0)} = 0.57$. The probabilities are determined as follows: the target flow from C to A is $b_{CA} + e_{CA} = 1.6$, while the number of matched cars from C to A is zero, leaving an (exact) deficit of $(1.6 - 0)$ in this direction; similarly, the (exact) deficit of the C-to-B direction is $(1.1+1.1-1) = 1.2$; the deficit of the C-to-C direction is 0. We want the direction with higher deficit to receive relocation assignments with higher probability, so a natural way to do so is as above. (If it turns out that there is zero deficit in each direction originating from C, we just randomly and independently relocate cars to all directions with equal probabilities.) The relocation in FA-mdR policy is a "loose mimicking" of the deterministic solution for two reasons: firstly, only directions with positive deficits receives assignments with positive probability, which makes sense if the total number of cars to be relocated are not sufficient to achieve the target deterministic flow but arguably does not make as much sense if there are more than sufficient cars to achieve the target flow; secondly, under independent assignments it could be the case that all the 3 remaining cars are assigned to the same direction, or in other words, the ex post realization of relocation could be far from the target flow.

- FA-rR: each of the 3 remaining cars is independently and randomly assigned to any direction.

- FA-woR: all three remaining cars simply stay at location C.

Table 2.1. Demand Function Parameters and Optimal Deterministic Solution of a Base Traffic Network

This table displays parameters and optimal deterministic solution for a base traffic network. The traffic network contains 3 locations, and the fleet of the ride-sharing platform contains 3 cars. The ride requests from location i to j has a Poisson distribution with density $\lambda_{ij}(p_{ij}) = \Lambda_{ij}e^{-\alpha_{ij}p_{ij}}$, in which Λ_{ij} captures the arrival rate of riders, and $-\alpha_{ij}$ captures the price sensitivity of riders. $((p_{ij}), (b_{ij}), (e_{ij}))$ constitutes the solution to the corresponding deterministic problem, in which p_{ij} is price, b_{ij} is matching rate, and e_{ij} is relocation rate.

Routes		Parameters		Opt. Det. Sol.		
i	j	Λ_{ij}	α_{ij}	p_{ij}	b_{ij}	e_{ij}
1	1	0.0	1.0	0.0	0.0	0.0
1	2	1.0	2.0	0.585	0.310	0.0
1	3	2.0	3.0	0.504	0.441	0.0
2	1	0.7	0.8	1.335	0.240	0.0
2	2	0.0	1.0	0.0	0.0	0.0
2	3	3.0	2.0	0.671	0.784	0.0
3	1	0.5	1.0	1.000	0.184	0.326
3	2	1.0	2.0	0.500	0.368	0.347
3	3	0.0	1.0	0.0	0.0	0.0

- G-SSS: exactly one car will be relocated to A, B, C respectively - see the previous subsection for a detailed implementation of G-SSS policy.

2.4.3 Performances of SSS and FA Policies: Simulation Evidence

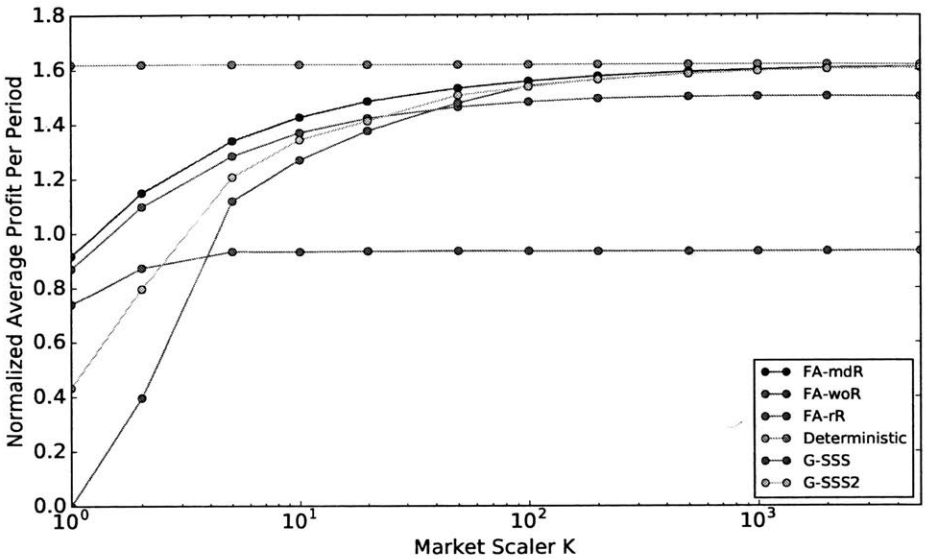
Now I present some simulation evidence which confirms that G-SSS heuristics are asymptotically optimal and suggests that FA-mdR policies are also asymptotically optimal, while FA-woR and FA-rR are not.

In the simulation I assume that there are 3 locations in the base traffic network, and it takes 1 period to travel between any two locations. In the base case, the fleet of the ride-sharing platform has 3 cars, and in each period the number of requests from i to j has Poisson distribution with Poisson density $\lambda_{ij}(p_{ij}) = \Lambda_{ij}e^{-\alpha_{ij}p_{ij}}$, where $\Lambda_{ij} \geq 0$ and $\alpha_{ij} \geq 0$ are origin-destination pair-specific parameters. The parameter specifications and the solution to the corresponding deterministic problem are shown in Table 2.1. As can be seen, the optimal solution to the corresponding deterministic problem is unique.

I scale up (by a factor of K) the number of vehicles and demand intensity in the base traffic network to investigate the asymptotic properties of the heuristic policies. In each

Figure 2-3. Convergence of Average Payoff for Different Policies in a Simulated Traffic Network

This figure shows the convergency of average payoffs for five heuristic policies and the deterministic benchmark. Check subsection 2.4.1 for the definition of the G-SSS strategy and subsection 2.4.2 for the definitions of the FA family of strategies. G-SSS2 is a modified version of G-SSS in which the dispatching target is defined as the integer ceil rather than floor of the sum of matching rate ($b^{\star}_{ij}$) and relocation rate ($e^{\star}_{ij}$) in the solution to the corresponding deterministic problem. For each strategy, I simulate 10 paths, and for each path, I simulate the traffic network for 5000 periods.



scaled traffic network, for each heuristic I generate $m = 10$ simulation path, with each path lasting for $T = 5000$ periods. For each path, starting with some initial state x_0, I record the realized payoff in each period and calculate the average payoff of those T periods. I then take the average performance of the m paths (i.e. the average of the m average payoff) to form an estimation of the long-run average performance $J_{\pi}(x_0)$. [16]

The simulation results for average payoff for each policy are shown in Figure 2-3. From the plot we can see that as the scale of the economy increases, the average payoffs of the G-SSS and FA-mdR policies (normalized by a factor of K) converge to the theoretical upper bound. This confirms our theoretical result about asymptotic optimality of GSSS policy, and

[16]Strictly speaking, each path gives a realization of the average payoff in the first T periods (starting at state x_0), while I want to know the expected long-run average revenue (i.e. $T \to \infty$). Intuitively, when T is large enough, the expected average payoff in the first T periods should be very close to the expected long-run average payoff. Furthermore, when T is large, the variance of the realized average payoff of the first T periods should also be small; I can form an estimation of this variance by looking at the sample standard deviation of the m paths. The bottom line is that via simulation I can only form an estimation of the long-run perfomance of each heuristic, but such estimation should be accurate enough for our purpose.

further suggests that FA-mdR policy is also asymptotically optimal. In addition, it seems that FA-mdR policy has better performance in small markets compared to GSSS policy. In the meanwhile, we see that FA-woR and FA-rR heuristics are not asymptotically optimal, which highlights the importance of relocation decisions.

In addition to the G-SSS policy I defined in subsection 2.4.1, I also checked the simulation results for a modified version of G-SSS, called G-SSS2. The difference is that in G-SSS, I define the i-to-j dispatching target to be $d^\star_{ij} = \left\lfloor b^\star_{ij} + e^\star_{ij} \right\rfloor$, while in G-SSS2, I define the dispatching target to be $d^\star_{ij} = \left\lceil b^\star_{ij} + e^\star_{ij} \right\rceil$ (where $\lceil x \rceil$ is the smallest integer bigger than or equal to x). It turns out that this small modification significantly improves the performance of G-SSS policy in small markets, without affecting its asymptotic optimality. Intuitively, this is because the larger dispatching targets (thus larger quotas) under G-SSS2 policy leaves fewer idle cars running from location i to itself.[17]

2.5 Small Market Performance

As shown in the previous section, the performance of FA-mdR policy converges to the theoretical upper bound in the long run, and FA-mdR also has better small market performance than other heuristic candidate policies. In this section, I compare the performance of FA-mdR with the exact optimal policy in a small traffic network.

2.5.1 Computational Method: Relative Value Iteration

Given the complexity of the problem, a closed-form exact solution is difficult to find, if not impossible. In fact, we would expect that even numerical solutions become difficult to obtain when the size of the problem is reasonably large, due to the extremely fast growth in computational complexity. For the purpose of demonstration and verification of some later results, I numerically solve the exact solution for a small scale traffic network.

The computational methods for average-payoff dynamic programming problems are more intricate compared to the discounted problems, and the validity of these methods may depend

[17] I think the asymptotic optimality of G-SSS2 policy can be proved by modifying the proof for the G-SSS policy, though I have not verified this yet.

Algorithm 1 Relative Value Iteration

$\epsilon > 0$ small, $k = 0$, x_0 is some fixed state (chosen arbitrarily)
for $x \in \{1, ..., S\}$ **do**
  $h^0(x) = 0$
end for
do
  for $x \in \{1, ..., S\}$ **do**
    $h^{k+1}(x) = \max_{u \in U(x)} \left\{ g(x,u) + \sum_{y=1}^{S} p_{xy}(u) h^k(y) \right\} - \max_{u \in U(x_0)} \left\{ g(x_0,u) + \sum_{y=1}^{S} p_{x_0 y}(u) h^k(y) \right\}$
  end for
while $\left\| h^{k+1} - h^k \right\| > \epsilon$
$h^\star = h^k$

on assumptions that relate to the structure of the underlying Markov chains. To find a solution to Bellman's equation (2.2.5), we consider the *relative value iteration* (VI) algorithm in Algorithm 1, in which we use Bellman equation to update the relative value function $h(\cdot)$ until it converges. The subsection below proves the theoretical validity of the relieve value iteration algorithm.

Theoretical Validity of Relative Value Iteration

If the sequence of relative value functions $\left\{h^k\right\}$ converges to a $h^\star$(i.e. $h^k(x) \to h^\star(x)$ for all x), then it is easy to see that

$$\lambda = \max_{u \in U(x_0)} \left\{ g(x_0, u) + \sum_{y=1}^{S} p_{x_0 y}(u) h^\star(y) \right\}$$

$$h = h^\star$$

solve Bellman's equation (2.2.5). It is known that the convergence of $\left\{h^k\right\}$ in the relative VI algorithm requires conditions stronger than existence of solution to Bellman's equation. When both the state space and the control space are finite, there exist sufficient conditions which guarantee the convergence of relative VI algorithm, but our problem does not satisfy any of those conditions. Nevertheless, here I show the convergence of the above relative VI algorithm under the assumption that $\tau_{ij} = 1$ for all locations i and j, i.e. the travel time needed between any two locations is 1.

Before stating the results we first introduce some shorthand notations. For any function

$h : X = \{1, ..., S\} \to \mathbb{R}$, we consider a functional mapping T defined as follows:

$$(Th)(x) = \max_{u \in U(x)} g(x, u) + \sum_{y \in \mathcal{X}} p_{xy}(u) h(y), \; x \in \{1, ..., S\}$$

Since $(Th)(\cdot)$ itself is a function defined on the state space X, we view T as a mapping that transforms the function h on X into the function Th on X; or alternatively, since X is finite, we may as well view h and Th as vectors in $\mathbb{R}^S$, thus view T as a mapping from $\mathbb{R}^S$ to $\mathbb{R}^S$. We denote by T^k the composition of the mapping T with itself k times.

The following proposition provides the convergence of the relative VI algorithm.

Proposition 1. *Assume that $\tau_{ij} = 1$ for all $i, j \in \{1, ..., L\}$. Fix a state x_0 and consider the relative VI algorithm*

$$h^{k+1}(x) = \left(Th^k\right)(x) - \left(Th^k\right)(x_0),$$

where h^0 is an arbitrary vector. Then the sequence $\left\{h^k\right\}$ converges to a vector $h^\star$ satisfying

$$(Th^\star)(x_0) + h^\star(x) = (Th^\star)(x),$$

so that by Fact 2.2.2, $(Th^\star)(x_0)$ is equal to the optimal average payoff.

The key to the proof of Proposition 1 is the following "contraction mapping" property of the mapping T. Define for all $h = (h(1), ..., h(S)) \in \mathbb{R}^S$,

$$H(h) = max\{h(x) \big| x = 1, ..., S\},$$

$$L(h) = min\{h(x) \big| x = 1, ..., S\},$$

$$||h|| = H(h) - L(h).$$

($||\cdot||$ is refered to as the *span seminorm.*)

Lemma 15. *Assume that $\tau_{ij} = 1$ for all $i, j \in \{1, ..., L\}$. There exists $\epsilon > 0$ such that for any h_1, $h_2 \in \mathbb{R}^S$, we have*

$$||Th_1 - Th_2|| \leq (1 - \epsilon) ||h_1 - h_2||$$

Proof. See Appendix 2.7.5. □

Based on Lemma 15, it is immediate to show that $\left\|h^{k+1} - h^k\right\| = \left\|Th^k - Th^{k-1}\right\| \leq (1-\epsilon)\left\|h^k - h^{k-1}\right\|$, and by noting that $h^k(x_0) = 0$ for all k, it is straightforward to see that for all x, $\left\{h^k(x)\right\}$ is a Cauchy sequence and therefore converges to a limit $h^*(x)$. Thus Proposition 1 is proved.

Implementation and Complexity Analysis

It should be pointed out that numerically solving the maximization problem in the iteration step of the relative VI algorithm is a non-trivial task given the infiniteness and the complexity of the control space. I outline the optimization steps below:

1. Given the current iteration of relative value h^k, state x, prices $p = (p_{ij})$, and realized demand ω, optimize over all feasible dispatching decisions to determined whether an available vehicle should accept some request (and if so, which request) or should be reallocated to some location (and if so, which location). Since the set of feasible decisions is finite (and is not a large number when the number of vehicles and the number of locations are small), we can find the optimal decision through grid search, i.e. calculating and comparing the total payoffs of all possible decisions and choose the best one. Let $\underline{g}\left(h^k, x, p, \omega\right)$ denote the optimal total payoff given $\left(h^k, x, p, \omega\right)$.

2. Calculate $\underline{f}\left(h^k, x, p\right) = E_\omega\left[\underline{g}\left(h^k, x, p, \omega\right)\right]$. Note that even though in theory there are infinitely many possible demand realizations of ω, those realizations with $\omega_{ij} > N$ in effect are no different from $\omega_{ij} = N$, so we can just assume that ω_{ij} follows truncated Poisson distribution in the first place. Therefore the expectation can be calculated as a finite sum of payoffs corresponding to all possible demand realizations weighted by corresponding probabilities.

3. Update the relative value function as $h^{k+1}(x) = \max_p \underline{f}\left(h^k, x, p\right)$.

The main difficulties involved are the following:

- The state space grows large quickly: Note that $S = L^N$, where S is the number of states, L is the total number of locations, and N is the total number of vehicles.

- The set of possible realizations of demand grows large even more quickly: $|\Omega| = (N+1)^{L^2}$, where $|\Omega|$ is the number of possible realizations of demand.

- $max_p \underline{f}(h^k, x, p)$ is a non-convex maximization problem, and global optimum is required according to the previous theory.

The second difficulty basically prevents me from solving a problem with more than two locations using the above method, because then it becomes infeasible to store the set of possible demands. Fortunately, with two locations and two vehicles, the first two issues are not a big problem, so that I can still hope to find the exact solution for such small-size problems numerically.

The third difficulty is a bit tricky to deal with and I still do not have a complete solution to it. I use software packages to solve $max_p \underline{f}(h^k, x, p)$ and the solving strategies are black boxes to me. I tried two kinds of optimazation methods, one that does local optimization, requires initial guess, and is deterministic in nature, and the other that attemps to find global optimum, does not require any input of initial guess, but is stochastic in nature (so may return different output each time for the same inputs).[18] I observe that global maximum are not always successfully found using either kind of method, even if the optimization is claimed to have terminated successfully. Actually, it is not easy to verify if some output is indeed a global optimum when I suspect that it is - the best I can do so far is to run global optimization algorithm several times, and run local optimization algorithm with different initial guesses, and if no "significantly" better values are found I come to "believe" that global maximum has been found.

My way to move forward is to accept the optimization output without figuring out whether global optimum is found in each step, use it in the iteration to update h^{k+1}, and see if such $\{h^k\}$ converges. I use the local optimization algorithms in the iteration steps, using some fixed initial guess p_0 in each step. What I observe so far is that, if p_0 is randomly chosen, then it is possible that $\{h^k\}$ fails to converge; but interestingly, if the initial guess p_0 is chosen to be the "monopoly prices" (to be explained later), it seems that $\{h^k\}$ typically converges to some h^*, and furthermore, in the optimization problem $max_p \underline{f}(h^*, x, p)$, with

[18] Global optimization method usually takes longer time to come up with a solution.

the initial guess of "monopoly prices" again, the local optimization algorithm seems to be able to end up with the global optimum (at least quite close to the global optimum).

A Numeric Example

Using relative value iteration algorithm outlined in the previous subsection, now I show a numeric solution of the optimal policy for a traffic network with two locations (A and B) and two cars. Suppose the demand functions for A-to-B and B-to-A markets are respectively as follows:

$$\lambda_{AB}(p_{AB}) = 3e^{-p_{AB}}$$

$$\lambda_{BA}(p_{BA}) = e^{-p_{BA}}$$

So the demand in the A-to-B market is higher. Recall that there are zero demands for A-to-A or B-to-B rides.

Table 2.2 shows the numerical solution to this problem. With two cars and two locations, according to the previous theoretical framework, there are four states in total; but since cars are anonymous, the two states in which each car occupies one location can be consolidated because they are essentially the same. Therefore we have three states, denoted for simplicity as $(0,2)$, $(1,1)$ and $(2,0)$, where the numbers are respectively the number of cars in location A and B.

The relative value of state $(0,2)$ is normalized to zero. We see that state $(2,0)$ has higher relative value than state $(0,2)$; this is consistent with our intuition since location A has higher demand. Also observe that for any origin-destination ride, the optimal price in state $(1,1)$ is slightly higher than the state in which both cars are located at the origin; this is mostly consistent with the static intuition that the supply of vehicles in state $(1,1)$ in either location is more limited, thus the corresponding monopoly price should be higher.

We can also see that the dispatching plan under the optimal policy has the property of flexible accommodation - that is, a ride request is never rejected when it can be accommodated. It remains an interesting theoretical question whether this flexible accommodation property holds for any optimal policy of any problem in general.[19] Another characteristic

[19]This flexible accommodation property of the optimal policy seems to hold for all the two-location prob-

Table 2.2. An Optimal Policy for A Two-location Two-car Economy

This table illustrates an optimal policy for a two-location two-car economy with $\Lambda_{AB} = 3$, $\Lambda_{BA} = 1$, $\alpha_{AB} = \alpha_{BA} = 1$. The optimal average payoff is 1.045. $*$ means that any price would be optimal (because there is no supply in that location anyway). As for allocations, e_{ij} denotes the number of vehicles relocated from i to j, and b_{ij} denotes the number of vehicles that are matched to a i-to-j request. Only non-zero allocation quantities are shown for compactness of demonstration. Note that the optimal policy may not be unique.

States (N_A, N_B)	Relative Values	Prices	Dispatching Plan: Demand Scenarios $(\omega_{AB}, \omega_{BA})$	Dispatching Plan: Allocations
(0,2)	0	$p_{AB} = *$, $p_{BA} = 1.080$	(0,0)	$e_{BA} = 1$, $e_{BB} = 1$
			(0,1)	$b_{BA} = 1$, $e_{BB} = 1$
			(0,2)	$b_{BA} = 2$
			(1,0)	$e_{BA} = 1$, $e_{BB} = 1$
			(1,1)	$b_{BA} = 1$, $e_{BB} = 1$
			(1,2)	$b_{BA} = 2$
			(2,0)	$e_{BA} = 1$, $e_{BB} = 1$
			(2,1)	$b_{BA} = 1$, $e_{BB} = 1$
			(2,2)	$b_{BA} = 2$
(1,1)	0.693	$p_{AB} = 1.452$, $p_{BA} = 1.173$	(0,0)	$e_{BA} = 1$, $e_{AB} = 1$
			(0,1)	$b_{BA} = 1$, $e_{AB} = 1$
			(0,2)	$b_{BA} = 1$, $e_{AB} = 1$
			(1,0)	$e_{BA} = 1$, $b_{AB} = 1$
			(1,1)	$b_{BA} = 1$, $b_{AB} = 1$
			(1,2)	$b_{BA} = 1$, $b_{AB} = 1$
			(2,0)	$e_{BA} = 1$, $b_{AB} = 1$
			(2,1)	$b_{BA} = 1$, $b_{AB} = 1$
			(2,2)	$b_{BA} = 1$, $b_{AB} = 1$
(2,0)	0.500	$p_{AB} = 1.415$, $p_{BA} = *$	(0,0)	$e_{AA} = 1$, $e_{AB} = 1$
			(0,1)	$e_{AA} = 1$, $e_{AB} = 1$
			(0,2)	$e_{AA} = 1$, $e_{AB} = 1$
			(1,0)	$e_{AA} = 1$, $b_{AB} = 1$
			(1,1)	$e_{AA} = 1$, $b_{AB} = 1$
			(1,2)	$e_{AA} = 1$, $b_{AB} = 1$
			(2,0)	$b_{AB} = 2$
			(2,1)	$b_{AB} = 2$
			(2,2)	$b_{AB} = 2$

Table 2.3. Heuristics v.s. Optimal Policy

Car Number	$N=2$	$N=2$	$N=3$	$N=4$	$N=2$	$N=3$	$N=4$
$[(\alpha_{AB}, \Lambda_{AB}), (\alpha_{BA}, \Lambda_{BA})]$	$[(1,3),(1,1)]$	$[(0.5,1),(1,1)]$			$[(1,6),(1,3)]$		
Deterministic Upperbound	1.466	1.104	1.104	1.104	2.890	3.183	3.301
Exact Optimal Policy	1.045	0.937	1.032	1.084	1.866	2.337	2.680
GSSS	0	0	0	0.923	0	1.746	1.618
GSSS2	0.501	0.461	0.923	1.002	1.827	2.204	2.572
FA-mdR	1.002	0.778	0.928	1.007	1.827	2.203	2.573
KOA	1.002	0.923	1.028	1.084	1.827	2.263	2.574
KOP	1.045	0.781	0.933	1.008	1.865	2.272	2.678
Solution Multiplicity	No	Yes	Yes	Yes	No	No	No

of the above optimal policy is that given the matching decision (b_{ij}), the relocation decision (e_{ij}) always maximizes the relative value of the next state. Intuitively this charateristic should hold in general, since conditional on the matching decision, the relocation decision only affects the continuation payoff and does not affect the current payoff, thus it should be chosen so that the next state has a relative value as high as possible.[20]

In the next subsection we look at the comparison between the optimal policy and its various heuristics.

2.5.2 Exact Optimal Policy v.s. Heuristics in Small Markets

In the example above, the average payoff per period under the optimal policy is 1.045. The first column of the table below shows the performances of the G-SSS, G-SSS2 and FA-mdR policy (corresponding to some optimal deterministic solution) for this case. We can see that FA-mdR policy still achieves 96% of the profit from the optimal policy, though GSSS2 and GSSS policies perform poorly in this small market.

We then twist the coefficients differently and study another two demand scenarios. From column 2 to 4, the total potential demand is relatively low, and the riders from A to B have a higher willingness to pay; while in column 5 to 7, we have a large demand in both directions. It turns out that two-location network with slightly more cars can still be handled by the

lems that I numerically solved, including cases with up to 4 cars.

[20]This second characteristic has some flavor of the GSSS policy - it strives to go to certain prefered state using relocation (though unlike GSSS policy, the optimal policy refuses to do so at the expense of unnecessary rejection). It is an interesting question whether the "steady states" in the GSSS policy has any connection with the state of highest relative value.

optimal policy algorithm, so we also study those two-location city with $N = 3$ and $N = 4$ cars.

We can see that in column 2, FA-mdR is still the better performer, though it achieves only 83% of profit compared to the optimal. This is likely due to the lower demand. In column 5, with a much higher demand, both GSSS2 and FA-mdR achieve 98% of the optimal profit. The reason GSSS has zero performance in all the $N = 2$ cases is due to a rounding problem which results in allocating zero quota of cars in either direction. As we can see, when we have more cars, GSSS would start to have a more normal performance; also, other heuristics would have better performance. And in all cases, FA-mdR seems to have a good performance.

To further break down the difference between FA-mdR policy and the optimal policy, we note that the two policies differ in two aspects: firstly, the optimal policy has state-contigent prices, while FA-mdR policy has fixed prices; secondly, though both exhibit flexible accommondation property in matching (as far as we see in those examples), they differ in their relocation decision. Therefore, I further look at two kinds of intermediate policies: one uses the fixed prices as in the optimal deterministic solution but adopts the dispatching plan in the optimal policy (thus named as KOA, "Keep Optimal Allocation"); another uses the state-contigent prices in the optimal policy (thus named as KOP, "Keep Optimal Prices") but uses the same relocation policy as in FA-mdR. The performance difference between KOA and the optimal policy, (or the difference between FA-mdR and KOP as well), should reflect the effect of mispricing, while the difference between KOP and the optimal policy, (or the difference between FA-mdR and KOA as well), should reflect the effect of suboptimal relocation plan.

Looking at the results in column 1, 2 and 5, we find that while in column 1 and 5 it seems that the main reason for suboptimality of FA-mdR compared to the optimal policy can be attributed to lack of state-contigent prices (because KOP policy basically achieves the full optimal profit), in column 2 it is the opposite: having optimal state-contigent prices but only suboptimal relocation plan does not help much with the performance, while having fixed prices but optimal relocation plan recovers the optimal profit. This indicates that FA-mdR policy can be suboptimal in different ways in different environments.

Finally, it should be pointed out that when the supply is abundant relative to demand

(as turns out to be the case in column 2 to 4), the solution to the deterministic problem is not unique, thus the heuristics are also not unique. In this above analysis we arbitrarily choose an optimal deterministic policy to base on when multiplicity exists; it remains an interesting question if this choice matters for the performance of heuristics, and if so, what kind of choice tends to be better.

2.6 Conclusion and Discussion

This paper studies the optimal pricing and allocation problem for a ride-sharing platform under the assumption of stationary stochastic demand and fixed supply, the objective of which being maximizing the average expected profit given an infinite horizon and no discounting. Since finding the exact optimal solution to this problem is computationally costly, we turn to study various heuristic policies inspired by the deterministic version of the problem.

We first show that the deterministic problem provides a payoff upperbound for our original problem, and based on this we prove that GSSS heuristic policy has asymptotically optimal performance. Simulation evidence verifies this result, and further suggests that FA-mdR heuristic policy is probably also asymptotically optimal, and seems to be superior to GSSS and GSSS2 policy in small markets. We then study the optimal policies in some tiny two-location economy where the computation of optimal policies is feasible, and compare the performances of various heuristics to it. We find that optimal policies seem to always accept a ride request whenever feasible (flexible accomodation), and that FA-mdR policy, which shares the flexible accomodation feature, again seems to have the best and most reliable performance in small markets.

One major result delivered by this paper is that a correct fixed-price policy does well asymptotically compared to the optimal dynamic pricing policy. This result is not that surprising once one notice that our underlying assumption about demand is time-invariant. One's intuition about the efficiency of dynamic pricing applies better in an environment where the fundamental demand (and supply) changes over time. From a practical point of finding good heuristic policies, though, it is unclear how our results in this paper will translate in a non-stationary demand environment where changes happen so frequently that

average payoff criterion is no longer appropriate, and would be an interesting direction for future research.

Another interesting direction is about uncertainty. We assume that though demand is stochastic, its probability is perfectly known, while in reality, this might be too strong an assumption. As pointed out in another paper (citation needed), dynamic pricing might be more robust in an environment with uncertainty. Thus it is interesting to know what kind of dynamic pricing and allocation policy would have robust performance in an environment with uncertainty.

2.7 Appendix

2.7.1 Proof of Lemma 13

Proof. For concreteness let x_0 be the state that all the N vehicles are in location 1 and are available. (Recall that the state of the system is described by the vector of the state of each vehicle, which in turn is described by where its next available location is and how many periods are left before it becomes available.) Now consider any state x. Note that for a vehicle in any state, if we want, it takes at most $\overline{T} \stackrel{def}{=} 2\max_{i,j} t_{ij}$ periods for it to arrive at location 1 - just let it finish its current ride, the remaining time of which is at most $\max_{i,j} t_{ij}$, and relocate it to location 1 which takes at most another $\max_{i,j} t_{i,j}$ periods. So starting from x_0, by taking some appropriate controls $\{\mu_0, ..., \mu_{T-1}\}$, the system can arrive at state x_0 deterministically in exactly $\overline{T}$ periods. The same thing is true if we start from x and want to deterministically arrive at x_0 - just note that we have in theory the option to ignore the ride request and focus on relocation, thus gaining deterministic control over the system, of course at the expense of generating zero revenue.

Without loss assume that $J^\star_\beta(x_0) \leq J^\star_\beta(x)$ (we can argue similarly for the other case). Since starting from x_0 we can take some policy π under which we first deterministically arrive at state x after $\overline{T}$ periods and then switch to the optimal policy starting at state x (in the case that optimal policy does not exist, we can always take a sequence of policies to

approximate the optimal revenue), we have

$$J_\beta^\star(x_0) \geq J_{\beta,\pi}(x_0) = \beta^{\overline{T}} J_\beta^\star(x)$$

Since per period expected revenue $g(x,u) \leq \overline{R}$ is bounded, we have $J_\beta^\star(x_0) \leq \frac{\overline{R}}{1-\beta}$, so we have

$$J_\beta^\star(x) - J_\beta^\star(x_0) \leq \left(1-\beta^{\overline{T}}\right) J_\beta^\star(x) \leq \frac{1-\beta^{\overline{T}}}{1-\beta}\overline{R} < \overline{T}\cdot\overline{R} \overset{def}{=} B$$

Note that B does not depend on β or x. Thus we show that $|h_\beta(x)|$ is uniformly bounded. □

2.7.2 Proof of Claims Made in the Proof of Theorem 6

Given a fixed non-randomized stationary policy μ and a fixed initial state x_0, let ρ_x be the long-run frequency of state $x \in \{1,...,S\}$ (i.e. $\rho^T = (\rho_1...\rho_S)$ is the x_0th row of the long-run frequency matrix $P_\mu^\star = \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1} P_\mu^t$). I now formally prove the following claims:

1. $b_{ij}(x) \leq \lambda_{ij}(p_{ij}(x))$, $\forall x \in X = \{1,...,S\}$, $\forall i,j \in \{1,2,...,L\}$;
2. $\sum_x \rho_x \sum_j (b_{ij}(x) + e_{ij}(x)) = \sum_x \rho_x \sum_k (b_{ki}(x) + e_{ki}(x))$, $\forall i \in \{1,2,...,L\}$;
3. $\sum_x \rho_x \left(\sum_i \sum_j (b_{ij}(x) + e_{ij}(x)) t_{ij}\right) = N$.

Proof of Claim 1. This is straightforward: for any $x \in X = \{1,...,S\}$, $i,j \in \{1,...,L\}$, we have $b_{ij}(x) = E\left[b_{ij}(x,\omega)\middle|x\right] \leq E\left[\omega_{ij}\middle|x\right] = \lambda_{ij}(p_{ij}(x))$, where the first equality comes from the definition of $b_{ij}(x)$ which is introduced in Section 3.2 at the beginning of the proof of Theorem 6(note that we sometimes suppress the dependence of everything on μ because μ is given and fixed), the inequality comes from the feasibility of policy μ which requires that $b_{ij}(x,\omega) \leq \omega_{ij}$ (i.e. that number of busy cars must not exceed the realized demand), and the last equality holds because $\omega_{ij}\middle|x,\mu \sim Poisson(\lambda_{ij}(p_{ij}(x)))$.

Proof of Claim 2. Fix some location i. We want to prove the following:

$$(\rho_1 \ ... \ \rho_S)\begin{pmatrix} \sum_j (b_{ij}(1) + e_{ij}(1)) \\ \vdots \\ \sum_j (b_{ij}(S) + e_{ij}(S)) \end{pmatrix} = (\rho_1 \ ... \ \rho_S)\begin{pmatrix} \sum_k (b_{ki}(1) + e_{ki}(1)) \\ \vdots \\ \sum_k (b_{ki}(S) + e_{ki}(S)) \end{pmatrix}$$

Firstly, note that $(\rho_1 \ldots \rho_S)$ is the x_0th row of $P_\mu^\star$, and $P_\mu^\star$ satisfies that $P_\mu^\star P_\mu = P_\mu^\star$(see Proposition 5.1.1 of Berstekas' DP book), therefore $(\rho_1 \ldots \rho_S) P_\mu^k = (\rho_1 \ldots \rho_S)$ for any $k = 0, 1, 2, \ldots$. From now on we suppress the dependence on μ since μ is fixed throughout the rest of the discussion, so we only write P and $P^\star$. Let P_{xy} denote the xyth element of matrix P (i.e. the element in row x column y), and let P_{xy}^k be the xyth element of matrix P^k. Note that we have $P_{xy}^k = \sum_{z \in X} P_{xz}^t P_{zy}^{k-t}$ for any $0 \leq t \leq k$ (which says that starting with state x, the probability of arriving at state y in k period is equal to the summation of probabilities that the intermediate state at period $t \leq k$ is state z, over all state z).

Secondly, let $K_i \stackrel{def}{=} \max_{k \in \{1,\ldots,L\}} t_{ki}$, that is, it takes at most K_i periods to travel from any location k in the network to location i. We then partition the set of locations $\{1, \ldots, L\}$ into the following subsets:

$$i(\tau) = \{k \in \{1, \ldots, L\} \,|t_{ki} = \tau\}, \ \tau = 1, \ldots, K_i$$

that is, $i(\tau)$ is the subset of locations from which it takes τ periods to travel to i.

Thus we can write

$$\begin{aligned}
& (\rho_1 \ldots \rho_S) \begin{pmatrix} \sum_j (b_{ij}(1) + e_{ij}(1)) \\ \vdots \\ \sum_j (b_{ij}(S) + e_{ij}(S)) \end{pmatrix} \\
= \ & (\rho_1 \ldots \rho_S) P^{K_i} \begin{pmatrix} \sum_j (b_{ij}(1) + e_{ij}(1)) \\ \vdots \\ \sum_j (b_{ij}(S) + e_{ij}(S)) \end{pmatrix} \\
= \ & (\rho_1 \ldots \rho_S) \begin{pmatrix} \sum_x P_{1x}^{K_i} \sum_j (b_{ij}(x) + e_{ij}(x)) \\ \vdots \\ \sum_x P_{Sx}^{K_i} \sum_j (b_{ij}(x) + e_{ij}(x)) \end{pmatrix} \qquad (2.7.1)
\end{aligned}$$

where the first equality comes from our previous conclusion that $(\rho_1 \ldots \rho_S) P_\mu^k = (\rho_1 \ldots \rho_S)$ for any $k = 0, 1, 2, \ldots$, and the second equality is simply executing matrix multiplication.

Now we study the expression $\sum_x P_{1x}^{K_i} \sum_j (b_{ij}(x) + e_{ij}(x))$. We have

$$\begin{aligned}
& \sum_x P_{1x}^{K_i} \sum_j (b_{ij}(x) + e_{ij}(x)) \\
= \; & E\left[\sum_j (b_{ij}(x_{K_i}, \omega_{K_i}) + e_{ij}(x_{K_i}, \omega_{K_i})) \middle| x_0 = 1, \mu\right] \\
= \; & E\left[E\left[\sum_j (b_{ij}(x_{K_i}, \omega_{K_i}) + e_{ij}(x_{K_i}, \omega_{K_i})) \middle| x_0 = 1, x_1, ..., x_{K_i}, \mu\right] \middle| x_0 = 1, \mu\right]
\end{aligned}$$

where the first equality uses the observation that $P_{1x}^{K_i} = Prob\left(x_{K_i} = x \middle| x_0 = 1, \mu\right)$, and the second equality comes from law of iterated expectation.

Note that $\sum_j (b_{ij}(x_{K_i}, \omega_{K_i}) + e_{ij}(x_{K_i}, \omega_{K_i})) \middle| x_0 = 1, x_1, ..., x_{K_i}, \mu$ is the expected number of vehicles that are sent out of location i in period K_i conditioning on knowing that the initial state is $x_0 = 1$, that the subsequent states are x_1, ..., x_{K_i} respectively, and that the policy is the stationary policy μ. On the one hand, once we know x_{K_i}, we know from policy feasibility requirement that $\sum_j (b_{ij}(x_{K_i}, \omega_{K_i}) + e_{ij}(x_{K_i}, \omega_{K_i})) = N^{K_i}(i, 0)$ where $N^{K_i}(i, 0)$ is the number of cars in car state $(i, 0)$ in system state x_{K_i}, and this holds regardless of the demand realization ω_{K_i}. On the other hand, conditioning on knowing $x_0 = 1, x_1, ..., x_{K_i - 1}, \mu$, by applying the state transition dynamics (2.2.2) iteratively and noticing that $N^{K_i - K_i}(i, K_i) = 0$ (because a car that heads from any location j to location i is at most $K_i - 1$ periods away from arriving at location i at the beginning of any period), we have

$$\begin{aligned}
N^{K_i}(i, 0) &= N^{K_i - 1}(i, 1) + \sum_{k \in \{1, ..., L\}: t_{ki} = 1} \left(b_{ki}^{K_i - 1}(x_{K_i - 1}, \omega_{K_i - 1}) + e_{ki}^{K_i - 1}(x_{K_i - 1}, \omega_{K_i - 1})\right) \\
&= N^{K_i - 2}(i, 2) + \sum_{k \in \{1, ..., L\}: t_{ki} = 2} \left(b_{ki}^{K_i - 2}(x_{K_i - 1}, \omega_{K_i - 1}) + e_{ki}^{K_i - 2}(x_{K_i - 1}, \omega_{K_i - 1})\right) \\
&\quad + \sum_{k \in \{1, ..., L\}: t_{ki} = 1} \left(b_{ki}^{K_i - 1}(x_{K_i - 1}, \omega_{K_i - 1}) + e_{ki}^{K_i - 1}(x_{K_i - 1}, \omega_{K_i - 1})\right) \\
&= \cdots \\
&= \sum_{\tau \in \{1, ..., K_i\}} \sum_{k \in i(\tau)} (b_{ki}(x_{K_i - \tau}, \omega_{K_i - \tau}) + e_{ki}(x_{K_i - \tau}, \omega_{K_i - \tau}))
\end{aligned}$$

In other words, the vehicles that are currently available at location i must be sent to i from

some location k in some previous period. Combining the above two observations, we have

$$\begin{aligned}
&\sum_j \left(b_{ij}\left(x_{K_i},\omega_{K_i}\right)+e_{ij}\left(x_{K_i},\omega_{K_i}\right)\right)\Big| x_0=1,x_1,...,x_{K_i-1},\mu \\
=&\sum_{\tau\in\{1,...,K_i\}}\sum_{k\in i(\tau)}\left(b_{ki}\left(x_{K_i-\tau},\omega_{K_i-\tau}\right)+e_{ki}\left(x_{K_i-\tau},\omega_{K_i-\tau}\right)\right)\Big| x_0=1,x_1,...,x_{K_i-1},\mu
\end{aligned}$$

Therefore, taking expectations on both sides over the realizations of demands conditional on the known past states $x_0=1,x_1,...,x_{K_i}$, we have

$$\begin{aligned}
&E\left[\sum_j \left(b_{ij}\left(x_{K_i},\omega_{K_i}\right)+e_{ij}\left(x_{K_i},\omega_{K_i}\right)\right)\Big| x_0=1,x_1,...,x_{K_i},\mu\right] \\
=\ &E\left[\sum_{\tau\in\{1,...,K_i\}}\sum_{k\in i(\tau)}\left(b_{ki}\left(x_{K_i-\tau},\omega_{K_i-\tau}\right)+e_{ki}\left(x_{K_i-\tau},\omega_{K_i-\tau}\right)\right)\Big| x_0=1,x_1,...,x_{K_i-1},\mu\right]
\end{aligned}$$

Again taking expectations on both sides over possible realizations of historical states $x_1,...,x_{K_i}$, we can further write

$$\begin{aligned}
&E\left[E\left[\sum_j \left(b_{ij}\left(x_{K_i},\omega_{K_i}\right)+e_{ij}\left(x_{K_i},\omega_{K_i}\right)\right)\Big| x_0=1,x_1,...,x_{K_i},\mu\right]\Big| x_0=1,\mu\right] \\
=\ &E\left[E\left[\sum_{\tau\in\{1,...,K_i\}}\sum_{k\in i(\tau)}\left(b_{ki}\left(x_{K_i-\tau},\omega_{K_i-\tau}\right)+e_{ki}\left(x_{K_i-\tau},\omega_{K_i-\tau}\right)\right)\Big| x_0=1,x_1,...,x_{K_i-1},\mu\right]\Big| x_0=1,\mu\right] \\
=\ &\sum_{(x_1,...,x_{K_i})} P_{1x_1}...P_{x_{K_i-1}x_{K_i}}E\left[\sum_\tau\sum_{k\in i(\tau)}\left(b_{ki}\left(x_{K_i-\tau},\omega_{K_i-\tau}\right)+e_{ki}\left(x_{K_i-\tau},\omega_{K_i-\tau}\right)\right)\Big| x_0=1,x_1,...,x_{K_i},\mu\right] \\
=\ &\sum_\tau\sum_{k\in i(\tau)}\left\{\sum_{x_1,...,x_{K_i}} P_{1x_1}...P_{x_{K_i-1}x_{K_i}}E\left[\left(b_{ki}\left(x_{K_i-\tau},\omega_{K_i-\tau}\right)+e_{ki}\left(x_{K_i-\tau},\omega_{K_i-\tau}\right)\right)\Big| x_0=1,x_1,...,x_{K_i},\mu\right]\right\} \\
=\ &\sum_\tau\sum_{k\in i(\tau)}E\left[b_{ki}\left(x_{K_i-\tau},\omega_{K_i-\tau}\right)+e_{ki}\left(x_{K_i-\tau},\omega_{K_i-\tau}\right)\Big| x_0=1,\mu\right] \\
=\ &\sum_\tau\sum_{k\in i(\tau)}\sum_{x_{K_i-\tau}} P^{K_i-\tau}_{1x_{K_i-\tau}}E\left[b_{ki}\left(x_{K_i-\tau},\omega_{K_i-\tau}\right)+e_{ki}\left(x_{K_i-\tau},\omega_{K_i-\tau}\right)\Big| x_0=1,x_{K_i-\tau},\mu\right] \\
=\ &\sum_\tau\sum_{k\in i(\tau)}\sum_{x_{K_i-\tau}} P^{K_i-\tau}_{1x_{K_i-\tau}}\left(b_{ki}\left(x_{K_i-\tau}\right)+e_{ki}\left(x_{K_i-\tau}\right)\right)
\end{aligned}$$

where the second equality comes from the definition of conditional expectation and the fact

that conditional on $x_0 = 1$ and policy μ, the probability that the states in period 1, 2, ..., K_i are respectively x_1, x_2, ..., $x_{K_i}$is $P_{1x_1}P_{x_1x_2}...P_{x_{K_i-1}x_{K_i}}$; the third equality uses the linearity of expectation operator and further rearranges terms; the fourth equality uses law of iterated expectation again; the fifth equality further uses law of iterated expectation by conditioning on $x_{K_i-\tau}$; and the last equality holds because by definition $b_{ki}(x_{K_i-\tau}) = E\left[b_{ki}(x_{K_i-\tau}, \omega_{K_i-\tau}) \middle| x_{K_i-\tau}, \mu\right]$, and furthermore, conditional on $x_{K_i-\tau}$, the realization of $\omega_{K_i-\tau}$ is independent of any previous state, therefore $E\left[b_{ki}(x_{K_i-\tau}, \omega_{K_i-\tau}) \middle| x_{K_i-\tau}, \mu\right] = E\left[b_{ki}(x_{K_i-\tau}, \omega_{K_i-\tau}) \middle| x_0 = 1, x_{K_i-\tau}, \mu\right]$ (In a similar way, we can show that $e_{ki}(x_{K_i-\tau}) = E\left[e_{ki}(x_{K_i-\tau}, \omega_{K_i-\tau}) \middle| x_0 = 1, x_{K_i-\tau}, \mu\right]$).

Therefore, we just show that

$$\sum_x P_{1x}^{K_i} \sum_j (b_{ij}(x) + e_{ij}(x)) = \sum_\tau \sum_{k\in i(\tau)} \sum_{x_{K_i-\tau}} P_{1x_{K_i-\tau}}^{K_i-\tau} (b_{ki}(x_{K_i-\tau}) + e_{ki}(x_{K_i-\tau})).$$

Similarly for state $y \in \{1, ..., S\}$,

$$\sum_x P_{yx}^{K_i} \sum_j (b_{ij}(x) + e_{ij}(x)) = \sum_\tau \sum_{k\in i(\tau)} \sum_{x_{K_i-\tau}} P_{yx_{K_i-\tau}}^{K_i-\tau} (b_{ki}(x_{K_i-\tau}) + e_{ki}(x_{K_i-\tau})).$$

Continue rewriting (2.7.1), we have

$$(\rho_1 \ ... \ \rho_S) \begin{pmatrix} \sum_x P_{1x}^{K_i} \sum_j (b_{ij}(x) + e_{ij}(x)) \\ \vdots \\ \sum_x P_{Sx}^{K_i} \sum_j (b_{ij}(x) + e_{ij}(x)) \end{pmatrix}$$

$$= (\rho_1 \ ... \ \rho_S) \begin{pmatrix} \sum_\tau \sum_{k\in i(\tau)} \sum_{x_{K_i-\tau}} P_{1x_{K_i-\tau}}^{K_i-\tau} (b_{ki}(x_{K_i-\tau}) + e_{ki}(x_{K_i-\tau})) \\ \vdots \\ \sum_\tau \sum_{k\in i(\tau)} \sum_{x_{K_i-\tau}} P_{yx_{K_i-\tau}}^{K_i-\tau} (b_{ki}(x_{K_i-\tau}) + e_{ki}(x_{K_i-\tau})) \end{pmatrix}$$

$$= \sum_\tau \sum_{k\in i(\tau)} (\rho_1 \ ... \ \rho_S) P^{K_i-\tau} \begin{pmatrix} (b_{ki}(1) + e_{ki}(1)) \\ \vdots \\ (b_{ki}(S) + e_{ki}(S)) \end{pmatrix}$$

$$= \sum_{\tau} \sum_{k \in i(\tau)} (\rho_1 \ldots \rho_S) \begin{pmatrix} (b_{ki}(1) + e_{ki}(1)) \\ \vdots \\ (b_{ki}(S) + e_{ki}(S)) \end{pmatrix}$$

$$= (\rho_1 \ldots \rho_S) \begin{pmatrix} \sum_k (b_{ki}(1) + e_{ki}(1)) \\ \vdots \\ \sum_k (b_{ki}(S) + e_{ki}(S)) \end{pmatrix}$$

where the first equality simply applies our previous result; the second equality can be verified by rearranging terms and checking matrix multiplication; the third equality uses our previous result that $(\rho_1 \ldots \rho_S))P^t = (\rho_1 \ldots \rho_S)$ for any t; the last equality is again simple rearrangement of terms.

Now we have what we want. Done.

Proof of Claim 3. We want to prove that

$$(\rho_1 \ldots \rho_S) \begin{pmatrix} \sum_{i,j} (b_{ij}(1) + e_{ij}(1)) t_{ij} \\ \vdots \\ \sum_{i,j} (b_{ij}(S) + e_{ij}(S)) t_{ij} \end{pmatrix} = N$$

Firstly, note that since $P^* P^* = P^*$ (again see Proposition 5.1.1 of Berstekas' DP book), we have $(\rho_1 \ldots \rho_S) P^* = (\rho_1 \ldots \rho_S)$. Thus we can write

$$(\rho_1 \ldots \rho_S) \begin{pmatrix} \sum_{i,j} (b_{ij}(1) + e_{ij}(1)) t_{ij} \\ \vdots \\ \sum_{i,j} (b_{ij}(S) + e_{ij}(S)) t_{ij} \end{pmatrix}$$

$$= (\rho_1 \ldots \rho_S) P^* \begin{pmatrix} \sum_{i,j} (b_{ij}(1) + e_{ij}(1)) t_{ij} \\ \vdots \\ \sum_{i,j} (b_{ij}(S) + e_{ij}(S)) t_{ij} \end{pmatrix}$$

$$= (\rho_1 \ldots \rho_S) \left(\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} P^t \right) \begin{pmatrix} \sum_{i,j} (b_{ij}(1) + e_{ij}(1)) t_{ij} \\ \vdots \\ \sum_{i,j} (b_{ij}(S) + e_{ij}(S)) t_{ij} \end{pmatrix}$$

$$= (\rho_1 \ldots \rho_S) \lim_{T\to\infty} \frac{1}{T} \left(\sum_{t=0}^{T-1} P^t \begin{pmatrix} \sum_{i,j} (b_{ij}(1) + e_{ij}(1)) t_{ij} \\ \vdots \\ \sum_{i,j} (b_{ij}(S) + e_{ij}(S)) t_{ij} \end{pmatrix} \right)$$

where the first equality uses the previously mentioned fact that $(\rho_1 \ldots \rho_S) P^* = (\rho_1 \ldots \rho_S)$, the second equality uses the definition that $P^* = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} P^t$ (the limit exists because P is a transition probability matrix), and the last equality is a simple rearrangement of terms.

Let $P^t_{x.}$ denote the xth row of matrix P^t. We observe that

$$\begin{aligned} & P^t_{1.} \begin{pmatrix} \sum_{i,j} (b_{ij}(1) + e_{ij}(1)) t_{ij} \\ \vdots \\ \sum_{i,j} (b_{ij}(S) + e_{ij}(S)) t_{ij} \end{pmatrix} \\ = \; & E\left[\sum_{i,j} (b_{ij}(x_t, \omega_t) + e_{ij}(x_t, \omega_t)) t_{ij} | x_0 = 1, \mu \right] \end{aligned}$$

Therefore,

$$\begin{aligned} & \sum_{t=0}^{T-1} P^t \begin{pmatrix} \sum_{i,j} (b_{ij}(1) + e_{ij}(1)) t_{ij} \\ \vdots \\ \sum_{i,j} (b_{ij}(S) + e_{ij}(S)) t_{ij} \end{pmatrix} \\ = & \begin{pmatrix} E\left[\sum_{t=0}^{T-1} \sum_{i,j} (b_{ij}(x_t, \omega_t) + e_{ij}(x_t, \omega_t)) t_{ij} | x_0 = 1, \mu \right] \\ \vdots \\ E\left[\sum_{t=0}^{T-1} \sum_{i,j} (b_{ij}(x_t, \omega_t) + e_{ij}(x_t, \omega_t)) t_{ij} | x_0 = S, \mu \right] \end{pmatrix} \\ = & \sum_{i,j} \begin{pmatrix} E\left[\sum_{t=0}^{T-1} (b_{ij}(x_t, \omega_t) + e_{ij}(x_t, \omega_t)) t_{ij} \middle| x_0 = 1, \mu \right] \\ \vdots \\ E\left[\sum_{t=0}^{T-1} (b_{ij}(x_t, \omega_t) + e_{ij}(x_t, \omega_t)) t_{ij} \middle| x_0 = 1, \mu \right] \end{pmatrix} \end{aligned}$$

Now, observe that[21]

$$
\begin{aligned}
&\sum_{t=0}^{T-1}\left(\sum_{i,j}\left(b_{ij}\left(x_t,\omega_t\right)+e_{ij}\left(x_t,\omega_t\right)\right)t_{ij}\right)\\
=\;&\sum_{t=0}^{T-1}\sum_{i,j}\sum_{\tau=0}^{t_{ij}-1}\left(b_{ij}\left(x_t,\omega_t\right)+e_{ij}\left(x_t,\omega_t\right)\right)\\
&-\sum_{i.j}\sum_{\tau=1}^{t_{ij}-1}\left(b_{ij}\left(x_{-\tau},\omega_{-\tau}\right)+e_{ij}\left(x_{-\tau},\omega_{-\tau}\right)\right)\left(t_{ij}-\tau\right)\\
&+\sum_{i,j}\sum_{\tau=1}^{t_{ij}-1}\left(b_{ij}\left(x_{T-\tau},\omega_{T-\tau}\right)+e_{ij}\left(x_{T-\tau},\omega_{T-\tau}\right)\right)\left(t_{ij}-\tau\right).
\end{aligned}
$$

[21]This observation might take a bit effort to verify so we explain it more carefully. In order to prove it, note that we only need to show that for all $i,j\in\{1,2,...,L\}$

$$
\begin{aligned}
&\Sigma_{t=0}^{T-1}((b_{ij}(x_t,\omega_t)+e_{ij}(x_t,\omega_t))t_{ij})\\
=\;&\Sigma_{t=0}^{T-1}\Sigma_{\tau=0}^{t_{ij}-1}(b_{ij}(x_{t-\tau},\omega_{t-\tau})+e_{ij}(x_{t-\tau},\omega_{t-\tau}))\\
&-\Sigma_{\tau=1}^{t_{ij}-1}(b_{ij}(x_{-\tau},\omega_{-\tau})+e_{ij}(x_{-\tau},\omega_{-\tau}))(t_{ij}-\tau)\\
&+\Sigma_{\tau=1}^{t_{ij}-1}(b_{ij}(x_{T-\tau},\omega_{T-\tau})+e_{ij}(x_{T-\tau},\omega_{T-\tau}))(t_{ij}-\tau)
\end{aligned}
$$

This can be shown simply by breaking things down more carefully:

$$
\begin{aligned}
&\Sigma_{t=0}^{T-1}(b_{ij}(x_t,\omega_t)+e_{ij}(x_t,\omega_t))t_{ij}\\
=\;&t_{ij}\left\{[b_{ij}(x_0,\omega_0)+e_{ij}(x_0,\omega_0)]+...+[b_{ij}(x_{T-1},\omega_{T-1})+e_{ij}(x_{T-1},\omega_{T-1})]\right\}\\
=\;&\left\{\left([b_{ij}(x_{-(t_{ij}-1)},\omega_{-(t_{ij}-1)})+e_{ij}(x_{-(t_{ij}-1)},\omega_{-(t_{ij}-1)})]+...+\left[b_{ij}(x_0,\omega_0)+e_{ij}(x_0,\omega_0)\right]\right)\right.\\
&+...\\
&+\left([b_{ij}(x_0,\omega_0)+e_{ij}(x_0,\omega_0)]+...+\left[b_{ij}(x_{t_{ij}-1},\omega_{t_{ij}-1})+e_{ij}(x_{t_{ij}-1},\omega_{t_{ij}-1})\right]\right)\\
&+\left([b_{ij}(x_1,\omega_1)+e_{ij}(x_1,\omega_1)]+...+\left[b_{ij}(x_{t_{ij}},\omega_{t_{ij}})+e_{ij}(x_{t_{ij}},\omega_{t_{ij}})\right]\right)\\
&+...\\
&\left.+\left([b_{ij}(x_{T-t_{ij}},\omega_{T-t_{ij}})+e_{ij}(x_{T-t_{ij}},\omega_{T-t_{ij}})]+...+\left[b_{ij}(x_{T-1},\omega_{T-1})+e_{ij}(x_{T-1},\omega_{T-1})\right]\right)\right\}\\
&-\left\{[b_{ij}(x_{-(t_{ij}-1)},\omega_{-(t_{ij}-1)})+e_{ij}(x_{-(t_{ij}-1)},\omega_{-(t_{ij}-1)})]\cdot 1+...\right.\\
&\left.+[b_{ij}(x_{-1},\omega_{-1})+e_{ij}(x_{-1},\omega_{-1})]\cdot(t_{ij}-1)\right\}\\
&+\left\{[b_{ij}(x_{T-t_{ij}+1},\omega_{T-t_{ij}+1})+e_{ij}(x_{T-t_{ij}+1},\omega_{T-t_{ij}+1})]\cdot 1+...\right.\\
&\left.+[b_{ij}(x_{T-1},\omega_{T-1})+e_{ij}(x_{T-1},\omega_{T-1})]\cdot(t_{ij}-1)\right\}\\
=\;&\Sigma_{t=0}^{T-1}\Sigma_{\tau=0}^{t_{ij}-1}(b_{ij}(x_{t-\tau},\omega_{t-\tau})+e_{ij}(x_{t-\tau},\omega_{t-\tau}))\\
&-\Sigma_{\tau=1}^{t_{ij}-1}(b_{ij}(x_{-\tau},\omega_{-\tau})+e_{ij}(x_{-\tau},\omega_{-\tau}))(t_{ij}-\tau)\\
&+\Sigma_{\tau=1}^{t_{ij}-1}(b_{ij}(x_{T-\tau},\omega_{T-\tau})+e_{ij}(x_{T-\tau},\omega_{T-\tau}))(t_{ij}-\tau).
\end{aligned}
$$

Here we introduced periods $t < 0$ merely for convenience; note that we don't have to interpret $t = 0$ as the first period ever in history.

Furthermore, we observe that[22]

$$\sum_{i,j} \sum_{\tau=1}^{t_{ij}-1} \left(b_{ij}\left(x_{T-\tau}, \omega_{T-\tau}\right) + e_{ij}\left(x_{T-\tau}, \omega_{T-\tau}\right)\right) = N$$

[22]To see why this is true, first recall from (2.2.1) that

$$N = \Sigma_{(j,\tau)\in C} N^{t+1}(j,\tau) = \Sigma_j \Sigma_{\tau=0}^{max_i t_{ij}-1} N^{t+1}(j,\tau).$$

From system dynamics (2.2.2) we can write

$$\begin{aligned}
N^{t+1}(j,\tau) &= N^t(j,\tau+1) + \Sigma_{i:t_{ij}=\tau+1}(b_{ij}(x_t,\omega_t) + e_{ij}(x_t,\omega_t)) \\
&= N^{t-1}(j,\tau+2) + \Sigma_{i:t_{ij}=\tau+2}(b_{ij}(x_{t-1},\omega_{t-1}) + e_{ij}(x_{t-1},\omega_{t-1})) \\
&\quad + \Sigma_{i:t_{ij}=\tau+1}(b_{ij}(x_t,\omega_t) + e_{ij}(x_t,\omega_t)) \\
&= \ldots \\
&= \Sigma_{i\in j(\tau+1)}(b_{ij}(x_t,\omega_t) + e_{ij}(x_t,\omega_t)) + \ldots \\
&\quad + \Sigma_{i\in j(max_i t_{ij})}(b_{ij}(x_{t-(max_i t_{ij}-\tau-1)}, \omega_{t-(max_i t_{ij}-\tau-1)}) + e_{ij}(x_{t-(max_i t_{ij}-\tau-1)}, \omega_{t-(max_i t_{ij}-\tau-1)}))
\end{aligned}$$

where $j(\tau) = \{i \in \{1,2,\ldots,L\} : t_{ij} = \tau\}$. Thus we can write

$$\begin{aligned}
\Sigma_{\tau=0}^{max_i t_{ij}-1} N^{t+1}(j,\tau) &= N^{t+1}(j,0) + N^{t+1}(j,1) \ldots + N^{t+1}(j, max_i t_{ij} - 1) \\
&= \{\Sigma_{i\in j(1)}(b_{ij}(x_t,\omega_t) + e_{ij}(x_t,\omega_t)) + \ldots \\
&\quad + \Sigma_{i\in j(max_i t_{ij})}(b_{ij}(x_{t-(max_i t_{ij}-1)}, \omega_{t-(max_i t_{ij}-1)}) + e_{ij}(x_{t-(max_i t_{ij}-1)}, \omega_{t-(max_i t_{ij}-1)}))\} \\
&\quad + \{\Sigma_{i\in j(2)}(b_{ij}(x_t,\omega_t) + e_{ij}(x_t,\omega_t)) + \ldots \\
&\quad + \Sigma_{i\in j(max_i t_{ij})}(b_{ij}(x_{t-(max_i t_{ij}-2)}, \omega_{t-(max_i t_{ij}-2)}) + e_{ij}(x_{t-(max_i t_{ij}-2)}, \omega_{t-(max_i t_{ij}-2)}))\} \\
&\quad + \ldots \\
&\quad + \{\Sigma_{i\in j(max_i t_{ij})}(b_{ij}(x_t,\omega_t) + e_{ij}(x_t,\omega_t))\} \\
&= \Sigma_{i\in j(1)}(b_{ij}(x_t,\omega_t) + e_{ij}(x_t,\omega_t)) \\
&\quad + \Sigma_{i\in j(2)}[(b_{ij}(x_t,\omega_t) + e_{ij}(x_t,\omega_t)) + (b_{ij}(x_{t-1},\omega_{t-1}) + e_{ij}(x_{t-1},\omega_{t-1}))] \\
&\quad + \ldots \\
&\quad + \Sigma_{i\in j(max_i t_{ij})}[(b_{ij}(x_t,\omega_t) + e_{ij}(x_t,\omega_t)) + \ldots \\
&\quad + (b_{ij}(x_{t-(max_i t_{ij}-1)}, \omega_{t-(max_i t_{ij}-1)}) + e_{ij}(x_{t-(max_i t_{ij}-1)}, \omega_{t-(max_i t_{ij}-1)}))] \\
&= \Sigma_i \Sigma_{\tau=0}^{t_{ij}-1}(b_{ij}(x_{t-\tau},\omega_{t-\tau}) + e_{ij}(x_{t-\tau},\omega_{t-\tau}))
\end{aligned}$$

Therefore

$$\begin{aligned}
N &= \Sigma_j \Sigma_{\tau=0}^{max_i t_{ij}-1} N^{t+1}(j,\tau) \\
&= \Sigma_j \Sigma_i \Sigma_{\tau=0}^{t_{ij}-1}(b_{ij}(x_{t-\tau},\omega_{t-\tau}) + e_{ij}(x_{t-\tau},\omega_{t-\tau})).
\end{aligned}$$

Therefore, let $t_{max} = \max_{i,j} t_{ij}$, we have

$$\begin{aligned}
& \left| \sum_{t=0}^{T-1} \left(\sum_{i,j} \left(b_{ij}\left(x_t, \omega_t\right) + e_{ij}\left(x_t, \omega_t\right) \right) t_{ij} \right) - TN \right| \\
\leq & \left| \sum_{i,j} \sum_{\tau=1}^{t_{ij}-1} \left(b_{ij}\left(x_{-\tau}, \omega_{-\tau}\right) + e_{ij}\left(x_{-\tau}, \omega_{-\tau}\right) \right) \left(t_{ij} - \tau\right) \right| \\
& + \left| \sum_{i,j} \sum_{\tau=1}^{t_{ij}-1} \left(b_{ij}\left(x_{T-\tau}, \omega_{T-\tau}\right) + e_{ij}\left(x_{T-\tau}, \omega_{T-\tau}\right) \right) \left(t_{ij} - \tau\right) \right| \\
\leq & \sum_{i,j} \left| N t_{max}^2 \right| + \Sigma_{i,j} \left| N t_{max}^2 \right| \\
\leq & \; 2L^2 N t_{max}^2
\end{aligned}$$

Thus we have

$$\begin{aligned}
& \lim_{T\to\infty} \frac{1}{T} \left| \sum_{t=0}^{T-1} \left(\sum_{i,j} \left(b_{ij}\left(x_t, \omega_t\right) + e_{ij}\left(x_t, \omega_t\right) \right) t_{ij} \right) - TN \right| \\
\leq & \lim_{T\to\infty} \frac{1}{T} \left(2L^2 N t_{max}^2 \right) = 0
\end{aligned}$$

Therefore we have just shown that

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i,j} \left(b_{ij}\left(x_t, \omega_t\right) + e_{ij}\left(x_t, \omega_t\right) \right) t_{ij} = N$$

for all realizations of $(\omega_t)_{t=0,1,2,\ldots}$. Therefore,

$$
\begin{aligned}
& (\rho_1 \ldots \rho_S) \lim_{T\to\infty} \frac{1}{T} \left(\sum_{t=0}^{T-1} P^t \begin{pmatrix} \sum_{i,j} (b_{ij}(1) + e_{ij}(1)) t_{ij} \\ \vdots \\ \sum_{i,j} (b_{ij}(S) + e_{ij}(S)) t_{ij} \end{pmatrix} \right) \\
= \; & (\rho_1 \ldots \rho_S) \lim_{T\to\infty} \frac{1}{T} \begin{pmatrix} E\left[\sum_{t=0}^{T-1} \left(\sum_{i,j} (b_{ij}(x_t, \mu(x_t), \omega_t) + e_{ij}(x_t, \mu(x_t), \omega_t)) t_{ij} \right) \middle| x_0 = 1, \mu \right] \\ \vdots \\ E\left[\sum_{t=0}^{T-1} \left(\sum_{i,j} (b_{ij}(x_t, \mu(x_t), \omega_t) + e_{ij}(x_t, \mu(x_t), \omega_t)) t_{ij} \right) \middle| x_0 = S, \mu \right] \end{pmatrix} \\
= \; & (\rho_1 \ldots \rho_S) \begin{pmatrix} E\left[\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \left(\sum_{i,j} (b_{ij}(x_t, \mu(x_t), \omega_t) + e_{ij}(x_t, \mu(x_t), \omega_t)) t_{ij} \right) \middle| x_0 = 1, \mu \right] \\ \vdots \\ E\left[\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \left(\sum_{i,j} (b_{ij}(x_t, \mu(x_t), \omega_t) + e_{ij}(x_t, \mu(x_t), \omega_t)) t_{ij} \right) \middle| x_0 = S, \mu \right] \end{pmatrix} \\
= \; & (\rho_1 \ldots \rho_S) \begin{pmatrix} N \\ \vdots \\ N \end{pmatrix} \\
= \; & N
\end{aligned}
$$

Thus we complete the proof.

2.7.3 Proof of Lemma 14

Before presenting the proof, we first make some useful observations about G-SSS policy. Recall that $t_{ij} \equiv 1$ is assumed throughout the following discussion (unless otherwise specified).

Observation 1. Under G-SSS policy, if $N^t(i,0) \geq \sum_j \left(b^\star_{ij} + e^\star_{ij}\right) + B$ $(B \geq 1)$, then we must have:

(1) $\forall j \neq i$, $b^{t-1}_{ij}(\omega) + e^{t-1}_{ij}(\omega) = d^\star_{ij}$;

(2) $N^{t-1}(i,0) \geq \sum_j \left(b^\star_{ij} + e^\star_{ij}\right) + B - (L-1)$

Proof. Note that $N^t(i,0) = \sum_k \left(b^{t-1}_{ki}(\omega) + e^{t-1}_{ki}(\omega)\right)$ (this is the system dynamic for the case $t_{ij} \equiv 1$), and that for any $k \neq i$, $b^{t-1}_{ki}(\omega) + e^{t-1}_{ki}(\omega) \leq b^\star_{ki} + e^\star_{ki}$ (this is a feature of G-SSS policy);

further note that $\sum_{k\neq i}(b^\star_{ki}+e^\star_{ki}) = \sum_{j\neq i}\left(b^\star_{ij}+e^\star_{ij}\right)$. The above immediately suggests that $b^{t-1}_{ii}(\omega)+e^{t-1}_{ii}(\omega) \geq (b^\star_{ii}+e^\star_{ii})+B > d^\star_{ii}$. Since this says that there is excess supply of cars in period $t-1$ on the direction from i to i, it must be that there are no "deficits" from i to any other directions, i.e. for $\forall j \neq i$, $b^{t-1}_{ij}(\omega)+e^{t-1}_{ij}(\omega) = d^\star_{ij}$. Therefore, $N^{t-1}(i,0) = \sum_j\left(b^{t-1}_{ij}(\omega)+e^{t-1}_{ij}(\omega)\right) \geq \sum_{j\neq i}\left\lfloor b^\star_{ij}+e^\star_{ij}\right\rfloor + (b^\star_{ii}+e^\star_{ii})+B \geq \sum_j\left(b^\star_{ij}+e^\star_{ij}\right)+B-(L-1)$. □

Observation 2.

$$\left|\sum_{k\neq i} d^\star_{ki} - \sum_{j\neq i} d^\star_{ij}\right| < (L-1)$$

Proof. Note that $\sum_{k\neq i}(b^\star_{ki}+e^\star_{ki}) = \sum_{j\neq i}\left(b^\star_{ij}+e^\star_{ij}\right)$. Also, $(b^\star_{ki}+e^\star_{ki})-1 < d^\star_{ki} \leq b^\star_{ki}+e^\star_{ki}$, $\left(b^\star_{ij}+e^\star_{ij}\right)-1 < d^\star_{ij} \leq b^\star_{ij}+e^\star_{ij}$. This gives

$$\sum_{k\neq i}(b^\star_{ki}+e^\star_{ki})-(L-1) < \sum_{k\neq i} d^\star_{ki} \leq \sum_{k\neq i}(b^\star_{ki}+e^\star_{ki})$$

$$\sum_{j\neq i}\left(b^\star_{ij}+e^\star_{ij}\right)-(L-1) < \sum_{j\neq i} d^\star_{ij} \leq \sum_{j\neq i}\left(b^\star_{ij}+e^\star_{ij}\right)$$

which the observation immediately follows. □

Based on the above observations, we can further make the following important observation:

Observation 3. Suppose that $b^\star_{ij}+e^\star_{ij} > L-1$ for all i and $j \neq i$. Under G-SSS policy, if $N^t(i,0) \geq \sum_j\left(b^\star_{ij}+e^\star_{ij}\right)+(3L+1)(L-1)$ for some location i, then $N^{t-1}(i,0) > N^t(i,0)$.

Proof. Note that

$$N^{t-1}(i,0) = \sum_j\left(b^{t-1}_{ij}(\omega)+e^{t-1}_{ij}(\omega)\right) \tag{2.7.2}$$

$$N^t(i,0) = \sum_k\left(b^{t-1}_{ki}(\omega)+e^{t-1}_{ki}(\omega)\right) \tag{2.7.3}$$

where the first equation comes from the fact that G-SSS is an admissible policy, and the second equation is just the system dynamic. If $N^t(i,0) \geq \sum_j\left(b^\star_{ij}+e^\star_{ij}\right)+(3L+1)(L-1)$, then by Observation 1, we have

$$b^{t-1}_{ij}(\omega)+e^{t-1}_{ij}(\omega) = d^\star_{ij},\ \forall j \neq i \tag{2.7.4}$$

and

$$N^{t-1}(i,0) \geq \sum_j \left(b_{ij}^\star + e_{ij}^\star\right) + 3L(L-1) \tag{2.7.5}$$

Let $s_{ki}^{t-1} = d_{ki}^\star - \left(b_{ki}^{t-1}(\omega) + e_{ki}^{t-1}(\omega)\right)$ denote the "deficit" on the k-to-i direction in period $t-1$ (under G-SSS policy we know $0 \leq s_{ki}^{t-1} \leq d_{ki}^\star$). Then from (2.7.2), (2.7.3) and (2.7.4), we can write

$$\begin{aligned} N^{t-1}(i,0) - N^t(i,0) &= \sum_j \left(b_{ij}^{t-1}(\omega) + e_{ij}^{t-1}(\omega)\right) - \sum_k \left(b_{ki}^{t-1}(\omega) + e_{ki}^{t-1}(\omega)\right) \\ &= \sum_{j\neq i} d_{ij}^\star - \sum_{k\neq i}\left(d_{ki}^\star - s_{ki}^{t-1}\right) \\ &> \sum_{k\neq i} s_{ki}^{t-1} - (L-1) \end{aligned}$$

where the last inequality comes from Observation 2. Therefore to show that $N^{t-1}(i,0) > N^t(i,0)$ we only need to show that $\sum_{k\neq i} s_{ki}^{t-1} \geq (L-1)$, that is, the total "inflow deficit" from other locations to location i is sufficiently large in period $t-1$.

We now show that $\sum_{k\neq i} s_{ki}^{t-1} \geq (L-1)$. Note that

$$\sum_i N^{t-1}(i,0) = N = \sum_i \sum_j \left(b_{ij}^\star + e_{ij}^\star\right)$$

Together with (2.7.5), we have

$$\sum_{k\neq i} N^{t-1}(k,0) \leq \sum_{k\neq i}\sum_j \left(b_{kj}^\star + e_{kj}^\star\right) - 3L(L-1)$$

that is, in period $t-1$, in total there are "deficit" of available cars in other places. Therefore it is not surprising that such deficit under G-SSS policy translates into "inflow deficit" from other places to location i. Indeed, let $S_k^{t-1} = \sum_j \left(b_{kj}^\star + e_{kj}^\star\right) - N^{t-1}(k,0)$ denote the "car deficit" (which can be positive or negative) in location k, then on the one hand, we have

$$\sum_{k\neq i} S_k^{t-1} = \sum_{k\neq i}\left(\sum_j \left(b_{kj}^\star + e_{kj}^\star\right) - N^{t-1}(k,0)\right) \geq 3L(L-1) \tag{2.7.6}$$

On the other hand, we have

$$\sum_j \left(d_{kj}^\star - s_{kj}^{t-1}\right) = N^{t-1}(k,0) = \sum_j \left(b_{kj}^\star + e_{kj}^\star\right) - S_k^{t-1} \leq \sum_j \left(d_{kj}^\star + 1\right) - S_k^{t-1}$$

which implies that

$$\sum_j s_{kj}^{t-1} \geq S_k^{t-1} - L$$

It is easy to see that, under G-SSS policy which strives to balance the total deficit across different directions, the following must hold:

$$s_{ki}^{t-1} \geq \min\left\{d_{ki}^\star, \left\lfloor \frac{S_k^{t-1} - L}{L} \right\rfloor\right\} \geq \min\left\{d_{ki}^\star, \frac{S_k^{t-1} - 2L}{L}\right\}$$

Now, if there exists any location $k \neq i$ such that $d_{ki}^\star \leq \frac{S_k^{t-1}-2L}{L}$, then by the mere fact that $s_{ki}^{t-1} \geq d_{ki}^\star \geq L-1$ (the second inequality comes from the assumption that $b_{ij}^\star + e_{ij}^\star > L-1$), we already show that $\sum_{k\neq i} s_{ki}^{t-1} \geq (L-1)$. So suppose that for all location $k \neq i$, $d_{ki}^\star > \frac{S_k^{t-1}-2L}{L}$, so we further have

$$s_{ki}^{t-1} \geq \frac{S_k^{t-1} - 2L}{L}$$

Summing over all $k \neq i$,

$$\sum_{k\neq i} s_{ki}^{t-1} \geq \frac{\sum_{k\neq i} S_k^{t-1}}{L} - 2(L-1) \geq \frac{3L(L-1)}{L} - 2(L-1) = L-1$$

where the second inequality comes from (2.7.6). Thus we show that $\sum_{k\neq i} s_{ki}^{t-1} \geq L-1$, which completes the proof. □

We are now ready to prove Lemma 14. The key to the proof is Observation 3: if at period t, the number of cars at location i is "significantly more" than its "steady state level"($N^t(i,0) \geq \sum_j \left(b_{ij}^\star + e_{ij}^\star\right) + (3L+1)(L-1)$), then we can show that $N^{t-1}(i,0) > N^t(i,0)$, that is, the number of cars at location i must be even more at period $t-1$. This shows that for a given problem, such t must be bounded; it can only have limited "past days" because the total number of cars in the system is limited. In other words, such "far-from-steady-state" kind of state (i.e. $N(i,0) \geq \sum_j \left(b_{ij}^\star + e_{ij}^\star\right) + (3L+1)(L-1)$ for some

i) cannot have positive long-run probability, or alternatively speaking, there exists some $T > 0$ such that for all $t > T$, the system state $x_t = (N^t(i,0))$ must satisfy $N(i,0) < \sum_j \left(b_{ij}^\star + e_{ij}^\star\right) + (3L+1)(L-1)$. Thus we have proved the second inequality in the lemma.

As for the first inequality (which basically says that in the long run, the number of cars in each location cannot be much less than its steady state level), intuitively the proof is the following: If in period t, there is a large shortage of available cars in some location i, it must be the case that in period $t-1$, there were major shortages of car inflows from some locations (say k) to i, which in turn implies that in $t-1$, there were major shortages of car supply (compared to the steady state level) in location k. However we just show that in the long run, there won't be much excess of supply in any location, which dictates that the shortages of supply in any location must also be limited in any location in the long run since the total supply of cars among all the locations are just equal to the steady state total supply.

Formally, let $\overline{B} = (3L+1)(L-1)$, then we have just shown that there exists some $T > 0$ such that for all $t > T$, under G-SSS policy, $x_t = (N^t(i,0))_{(i,0)\in C}$ must satisfy $N^t(i,0) < \sum_j \left(b_{ij}^\star + e_{ij}^\star\right) + \overline{B}$ for all location i. Now suppose toward contradiction that $N^t(i,0) < \sum_k (b_{ki}^\star + e_{ki}^\star) - \underline{B}$ for some large t ($t > T$). We want to show that when $\underline{B}$ is sufficiently large (in fact $\underline{B} = 3L^2 - 1$ is sufficient), this must lead to contradiction.

As before let $s_{ki}^{t-1} = d_{ki}^\star - (b_{ki}^{t-1}(\omega) + e_{ki}^{t-1}(\omega))$ denote the "deficit" on the k-to-i direction in period $t-1$ (under G-SSS policy we know $0 \le s_{ki}^{t-1} \le d_{ki}^\star$). By system dynamics,

$$N^t(i,0) = \sum_k \left(b_{ki}^{t-1}(\omega) + e_{ki}^{t-1}(\omega)\right) = \Sigma_k \left(d_{ki}^\star - s_{ki}^{t-1}\right)$$

on the other hand, we assumed that

$$N^t(i,0) < \sum_k (b_{ki}^\star + e_{ki}^\star) - \underline{B} < \sum_k (d_{ki}^\star + 1) - \underline{B}$$

so we have

$$\sum_k s_{ki}^{t-1} > \underline{B} - L$$

Now note the following fact about G-SSS policy: if $s_{ki}^{t-1} \ge 1$, that is, if there is positive deficit of cars sent from k to i, there must be at least deficit of $s_{ki}^{t-1} - 1$ cars sent from k to

other directions in period $t-1$. Since G-SSS policy is admissible, we further have

$$\begin{aligned} N^{t-1}(k,0) &= \sum_j \left(b_{kj}^{t-1}(\omega) + e_{kj}^{t-1}(\omega)\right) \\ &= \sum_j \left(d_{kj}^{\star} - s_{kj}^{t-1}\right) \\ &\leq \sum_j d_{kj}^{\star} - L\left(s_{ki}^{t-1} - 1\right) \\ &< \sum_j \left(b_{kj}^{*} + e_{kj}^{*}\right) - L\left(s_{ki}^{t-1} - 1\right) \end{aligned}$$

Therefore

$$\begin{aligned} \sum_{k:s_{ki}^{t-1}\geq 1} N^{t-1}(k,0) &\leq \sum_{k:s_{ki}^{t-1}\geq 1} \sum_j \left(b_{kj}^{\star} + e_{kj}^{\star}\right) - L \sum_{k:s_{ki}^{t-1}\geq 1} s_{ki}^{t-1} + L^2 \\ &< \sum_{k:s_{ki}^{t-1}\geq 1} \sum_j \left(b_{kj}^{\star} + e_{kj}^{\star}\right) - L\left(\underline{B} - L\right) + L^2 \end{aligned} \tag{2.7.7}$$

where the second inequality uses the fact that

$$\sum_{k:s_{ki}^{t-1}\geq 1} s_{ki}^{t-1} = \sum_k s_{ki}^{t-1} > \underline{B} - L$$

On the other hand, for those k where in period $t-1$ the cars sent from k to i has zero deficit, merely by the fact that t is large enough, we know from our previous result that

$$N^{t-1}(k,0) < \sum_j \left(b_{kj}^{\star} + e_{kj}^{\star}\right) + \overline{B}$$

Thus

$$\sum_{k:s_{ki}^{t-1}=0} N^{t-1}(k,0) < \sum_{k:s_{ki}^{t-1}=0} \sum_j \left(b_{kj}^{\star} + e_{kj}^{\star}\right) + L\overline{B} \tag{2.7.8}$$

Combining (2.7.7) and (2.7.8) we have

$$\sum_k N^{t-1}(k,0) < \sum_k \sum_j \left(b_{kj}^{\star} + e_{kj}^{\star}\right) + L\overline{B} - L\left(\underline{B} - L\right) + L^2 \tag{2.7.9}$$

Noting that

$$\sum_k N^{t-1}(k,0) = N = \sum_k \sum_j \left(b^\star_{kj} + e^\star_{kj}\right)$$

we see that (2.7.9) requires that

$$L\overline{B} - L\left(\underline{B} - L\right) + L^2 > 0$$

which is equivalent to

$$\underline{B} < \overline{B} + 2L$$

Therefore as long as $\underline{B} \geq (3L+1)(L-1)+2L = 3L^2-1$, having $N^t(i,0) < \sum_k (b^\star_{ki} + e^\star_{ki}) - \underline{B}$ is impossible for t sufficiently large. This shows that in the long run, the system state must satisfy that $N^t(i,0) \geq \sum_k (b^\star_{ki} + e^\star_{ki}) - (3L^2-1)$. This completes the proof to Lemma 14.

2.7.4 Proof of Theorem 8

Proof. Again we only need to prove the second inequality. Fix any initial state x_0. Let $p^\star_{G-SSS}(x)$ be the long-run probability of state x under the G-SSS policy. Then

$$J^{G-SSS}(x_0) = \sum_{x \in X} p^\star_{G-SSS}(x)\, g_{G-SSS}(x)$$

where

$$g_{G-SSS}(x) = \sum_{i,j} p^\star_{ij} E_\omega \left[\omega_{ij} - \left(\omega_{ij} - q_{ij}(x)\right)^+\right]$$

By Lemma 14, we know that for any $x = (N(i,0))_{(i,0)\in C}$ such that $p^\star_{G-SSS}(x) > 0$, the following inequality holds:

$$N(i,0) \geq \sum_j \left(b^\star_{ij} + e^\star_{ij}\right) - 3L^2$$

therefore by the property of G-SSS policy we should further have

$$q_{ij}(x) \geq d^\star_{ij} - \frac{\sum_j \left(b^\star_{ij} + e^\star_{ij}\right) - N(i,0)}{L} - 1 > \left(b^\star_{ij} + e^\star_{ij} - 1\right) - 3L - 1 = b^\star_{ij} + e^\star_{ij} - (3L+2)$$

thus, combining with the Gallego (1992) result that for any random variable D with finite mean μ and finite standard deviation σ, and for any real number d,

$$E\left[(D-d)^{+}\right] \leq \frac{\sqrt{\sigma^2+(d-\mu)^2}-(d-\mu)}{2}$$

and applying this inequality and noting that $\omega_{ij} \sim Poisson\left(b_{ij}^{\star}\right)$ we obtain

$$\begin{aligned}
E_{\omega}\left[(\omega_{ij}-q_{ij}(x))^{+}\right] &\leq E_{\omega}\left[\left(\omega_{ij}-\left\{b_{ij}^{\star}+e_{ij}^{\star}-(3L+2)\right\}\right)^{+}\right] \\
&\leq \frac{\sqrt{b_{ij}^{\star}+\left(e_{ij}^{\star}-(3L+2)\right)^2}-\left(e_{ij}^{\star}-(3L+2)\right)}{2} \\
&\leq \frac{\sqrt{b_{ij}^{\star}}}{2}+\frac{\left|e_{ij}^{\star}-(3L+2)\right|-\left(e_{ij}^{\star}-(3L+2)\right)}{2} \\
&\leq \frac{\sqrt{b_{ij}^{\star}}}{2}+(3L+2)
\end{aligned}$$

Therefore,

$$\begin{aligned}
g_{G-SSS}(x) &= \sum_{i,j} p_{ij}^{\star} E_{\omega}\left[\omega_{ij}-(\omega_{ij}-q_{ij}(x))^{+}\right] \\
&\geq \sum_{i,j} p_{ij}^{\star} b_{ij}^{\star}-\sum_{i,j} p_{ij}^{\star}\left[\frac{\sqrt{b_{ij}^{\star}}}{2}+(3L+2)\right]
\end{aligned}$$

so we have

$$\begin{aligned}
J^{G-SSS}(x_0) &= \sum_{x\in X} p_{G-SSS}^{\star}(x)\, g_{G-SSS}(x) \\
&= \sum_{x\in X: p_{G-SSS}^{\star}(x)>0} p_{G-SSS}^{\star}(x)\, g_{G-SSS}(x) \\
&\geq \left(\sum_{x\in X: p_{G-SSS}^{\star}(x)>0} p_{G-SSS}^{\star}(x)\right)\left(\sum_{i,j} p_{ij}^{\star} b_{ij}^{\star}-\sum_{i,j} p_{ij}^{\star}\left[\frac{\sqrt{b_{ij}^{\star}}}{2}+(3L+2)\right]\right) \\
&= \sum_{i,j} p_{ij}^{\star} b_{ij}^{\star}-\sum_{i,j} p_{ij}^{\star}\left[\frac{\sqrt{b_{ij}^{\star}}}{2}+(3L+2)\right]
\end{aligned}$$

Dividing both sides by $J^D=\sum_{i,j} p_{ij}^{\star} b_{ij}^{\star}$ we immediately get the result in the theorem. □

2.7.5 Proof of Lemma 15

Proof. Recall that a control specifies the prices and the dispatching plan under all possible demand contingencies. A special demand contingency is that for all origin-destination pairs, the realized demands are zero ($\omega = (\omega_{ij}) = 0$). Let ϵ be the probability that all realized demands are zero when all the prices are set to be zero; since the Poisson intensities are bounded for all feasible prices (including zero), we know that $\epsilon > 0$, and for positive prices $\omega = 0$ happens with probabilities higher than ϵ.

Since $\tau_{ij} = 1$ for all origin-destination pairs, at the beginning of each period all the vehicles are available for matching, and starting at any state x, any state y can be reached in the next period with proper reallocation.

Let $\mu_1(x) \in \arg\max_{u \in U(x)} \left[g(x,u) + \sum_{y=1}^{S} p_{xy}(u) h_1(y)\right]$. If $\omega = 0$ is realized, according to $\mu_1(x)$, there is some plan of vehicle reallocation and next period the system would be in some state x_1. It must be that $x_1 \in \arg\max_y h_1(y)$, because otherwise suppose $x_1 \notin \arg\max_y h_1(y)$ but $x' \in \arg\max_y h_1(y)$, then modifying $\mu_1(x)$ by only changing the plan for contingency $\omega = 0$ from x_1 to x' does not affect expected current payoff (neither prices nor matching probabilities are changed), but strictly increases expected future payoff, thus contradicting the assumption that $\mu_1(x) \in \arg\max_{u \in U(x)} \left[g(x,u) + \sum_{y=1}^{S} p_{xy}(u) h_1(y)\right]$. Since $\omega = 0$ happens with probability of at least ϵ, $p_{xx_1}(\mu_1(x)) \geq \epsilon$.

Now suppose instead that the future payoff is h_2. Consider modifying $\mu_1(x)$ by only changing the plan for contingency $\omega = 0$ from x_1 to x_2 where $x_2 \in \arg\max_y h_2(y)$; doing so does not affect the expected current payoff but increases expected future payoff by at least $\epsilon(h_2(x_2) - h_2(x_2)) \geq 0$. Thus we can write:

$$g(x, \mu_1(x)) + \sum_{y=1}^{S} p_{xy}(\mu_k(x)) h_1(y) = (Th_1)(x)$$

$$g(x, \mu_1(x)) + \sum_{y=1}^{S} p_{xy}(\mu_1(x)) h_2(y) + \epsilon(h_2(x_2) - h_2(x_1)) \leq (Th_2)(x)$$

From these relations we obtain

$$\sum_{y=1}^{S} p_{xy}\left(\mu_1(x)\right)\left[h_2(y)-h_1(y)\right]+\epsilon\left(h_2(x_2)-h_2(x_1)\right)\leq(Th_2)(x)-(Th_1)(x)$$

Note that

$$\sum_{y=1}^{S} p_{xy}\left(\mu_1(x)\right)\left[h_2(y)-h_1(y)\right]\geq(1-\epsilon)L(h_2-h_1)+\epsilon\left(h_2(x_1)-h_1(x_1)\right)$$

Thus we have

$$(Th_2)(x)-(Th_1)(x)\geq(1-\epsilon)L(h_2-h_1)+\epsilon\left(H(h_2)-H(h_1)\right)$$

Therefore

$$L(h_2-h_1)\geq(1-\epsilon)L(h_2-h_1)+\epsilon\left(H(h_2)-H(h_1)\right)$$

Reversing the role of h_1 and h_2 we have

$$L(h_1-h_2)\geq(1-\epsilon)L(h_1-h_2)+\epsilon\left(H(h_1)-H(h_2)\right)$$

which can be rewritten as

$$H(h_2-h_1)\leq(1-\epsilon)H(h_2-h_1)+\epsilon\left(H(h_2)-H(h_1)\right)$$

Thus we have

$$\|h_2-h_1\|=H(h_2-h_1)-L(h_2-h_1)\leq(1-\epsilon)\left(H(h_2-h_1)-L(h_2-h_1)\right)=(1-\epsilon)\|h_2-h_1\|$$

□

Chapter 3

Collaborating: Discrete versus Continuous-Time Modeling

3.1 Introduction

Bonatti and Horner (2011) study an interesting model of dynamic moral hazard in teams where two symmetric agents collaborate on a project over some time with uncertainty about the state of the world. Agents cannot observe each other's effort, therefore cannot be sure if a persistent lack of success is due to a bad underlying state or the lack of effort from the other agent. The model predicts a unique symmetric equilibrium, in which the effort level is at first decreasing over time, then it jumps to the maximal level at some point when the deadline looms closely enough.

The Bonatti-Horner model adopts a continuous-time framework where agents are assumed to adjust their effort level instantaneously. The authors also mention a version of two-period model for illustrative purpose, but leave it unsolved. In this paper, I study several discrete-time versions of their model in detail. As it turns out, I find that depending on modeling assumptions, two-period versions of the model can give equilibrium results quite different from that in the continuous-time model. For example, while the continuous-time model predicts existence and uniqueness of symmetric equilibrium, its two-period versions can either have multiple symmetric equilibria or none. Also, not all equilibria have features similar to the one in the Bonatti-Horner model - most notably, two-period models sometimes

predict symmetric equilibria in which agents exert their maximal effort in the first period but none in the second, in contrast to the procrastination story told by the continuos-time model.

Studying related discrete-time models can provide an alternative perspective on the same question. Such a perspective is valuable because one may not always believe that agents are capable of processing information and reacting instantaneously. Discrete-time models can also help up understand the economic forces behind an argument.

The paper is organized as follows. Section 2.1 gives a review of the main result in the Bonatti-Horner model, which characterizes the unique symmetric equilibrium where effort levels are first decreasing over time then jump to maximal level when the deadline is close enough. Section 2.2 then discussed the general setup of the related discrete-time model which aims at properly "approximate" the continuous-time model in a way that allows coarser control for the agents - that is, instead of being able to adjust effor level instantaneously, agents can only fix their level of effort over a period of time. When it comes to approximating the success probability conditional on the good state, I notice that there are at least two different ways to make mathematically convenient assumptions. The version of two-period model mentioned in Bonatti and Horner (2011) essentially assumes what I call "additive technology": within a period, agents' efforts contribute to success probability additively, so the other agent's effort in the same period does not affect one's marginal contribution to success. Another version assumes "product technology": within a period, it is as if success happens to each agent independently (but double successes do not bring extra payoff), thus the other agent's effort reduces one's marginal contribution to success.

Section 3 analyzes two-period model with addtive technology (Section 3.1) and product technology (Section 3.2) respectively. In Section 3.1, under additive technology, Proposition 1 shows that symmetric equilibrium exists for all sets of parameter values we are interested in, but Proposition 2 shows that, multiple symmetric equilibria exist for an open set of parameters. In particular, under certain parameter conditions, we have a symmetric equilibrium in which agents exert maximal effort in the first period and none in the second, and this can happen regardless of the discounting factor.

In Section 3.2, with product technology, Proposition 3 shows the uniqueness of sym-

metric equilibrium, but Proposition 4 points out that symmetric equilibrium may not exist under some parameter conditions. Section 3.3 then discusses some observed difference in equilibrium results among different models.

In an effort to understand if discrete-time model may "approximate" the continuous-time model better with more periods, Section 4 looks at a special (and more tractable) case of the game where the state is known to be good. Without learning, multiple-period models can be easily solved by backward induction, thus we can have a clear view of how equilibria evolve as the discrete-time model "goes to its limit" when approximating the continuous-time model increasingly better. Proposition 5 and Proposition 6 characterizes the (generically) unique symmetric equilibrium for additive technology and product technology respectively. Proposition 6 shows a relatively straightforward picture of equilibrium approximation when the technology is productive: the equilibrium effort level predicted by continuous-time model and discrete-time model with product technology are very close as we have sufficient number of periods. However, Proposition 5 shows a messier picture of equilibrium approximation when the technology is additive: instead of the intermediate level of effort in the initial stage of the game as predicted by the continuous-time model, multi-period additive technology model predicts effort levels alternating between maximal level and zero initially.

Section 5 discusses directions for future research and concludes.

Related Literature. In terms of the nature of the problem it studies, this paper is related to several strands of literature including experimentation and innovation, free-riding in teams, and dynamic contributions to public goods—see Bonatti and Horner (2011) for a thorough review for papers before them. To briefly mention a few more recent papers, Guo (2016) study the optimal delegation rule in a dynamic relationship in which a principal delegates experimentation to an agent with private information about the nature of the experiment. Halac, Kartik, and Liu (2016) study a model of long-term contracting for experimentation in a principal-agent relationship, where adverse selection on the agent's ability, dynamic moral hazard, and private learning about project quality all play some role in structuring dynamic incentives. Georgiadis (2014) studies a dynamic problem in which a group of agents collaborate over time to complete a project, and finds that the optimal team

size decreases in the expected length of the project.

In terms of modeling technicality this paper is also related to papers studying the relationship between discrete and continuous time models and limits of discrete-time models as period lengths shorten. Sannikov (2008) studies the optimal contract in a continuous-time version of dynamic moral hazard problem in a principal-agent setting, the result of which shares similarities with some discrete-time model solutions, but the solution of the continuous-time model turns out to be much more tractable. DeMarzo and Sannikov (2006) study the optimal contract in a cash-flow diversion model using the methods from Sannikov (2008), and Biais, Mariotti, Plantin, and Rochet (2007) show that the contract of DeMarzo and Sannikov (2006) arises in the limit of discrete-time models as the agent's actions become more frequent.

3.2 Model Setup

In this paper we take a specific continuous-time model extensively studied by Bonatti and Horner (2011) and look at different versions of discrete-time models corresponding to it. Section 2.1 describes the continuous-time model and gives a review of some results in Bonatti and Horner (2011). Section 2.2 discusses different ways to transform the continuous-time model to discrete-time ones, and gives the setup of a general discrete-time game.

3.2.1 A Review of the Continuous-Time Model Results

There are two agents engaged in a common project, which has a probability of $\bar{p} < 1$ of being a good one (it is a bad project otherwise).[1] Agents continuously choose at which level to exert effort over the time horizon $[0, T]$, and the choice is restricted to the unit interval: $u_{i,t} \in [0, 1]$.[2] The instantaneous cost of effort is $c_i(u_{i,t}) = \alpha u_{i,t}$ for some constant $\alpha > 0$. It is assumed that $\alpha < \bar{p}$. The effort choice is and remains unobserved by the other agent.

If the project is good, a breakthrough or success occurs with instantaneous probability equal to $\Sigma_i u_{i,t}$, i.e. if agents were to exert a constant effort u_i over some interval of time, then

[1]The model can be easily generalized into one with n agents; in this paper we just focus on models with two agents for simplicity.

[2]The unit interval is just a normalization; one may alternatively choose to normalize the time horizon.

the delay until they find out that project is successful would be distributed exponentially over that time interval with parameter $\Sigma_i u_{i,t}$.[3]If the project is bad, success is impossible regardless of effort exerted.

The game ends if a breakthrough occurs, or if the deadline T is reached. A breakthrough is normalized to be worth 1 to each of the agents; agents reap no benefit without a breakthrough. Agents are impatient, and discount future benefits and costs at a common discount rate r.

This is a dynamic game with incomplete information. For a dynamic game the strategy of an agent should specify his action on every possible information set that he may reach. Since the action of the other agent is unobservable, each agent's strategy can be simply described as a measurable function $u_i : [0, T] \to [0, 1]$.

As time passes, the game either ends or keeps going, and in either case agents get updates about the state of the world: If the game ends (because of a breakthrough), then the state must be good; otherwise the state is still uncertain but the posterior belief must be (weakly) more pessimistic.

We mainly focus on symmetric equilibrium. Bonatti and Horner (2011) have the following result for this game: [4]

Theorem 9. *Suppose $\frac{1}{r} \geq \alpha^{-1} - \bar{p}^{-1} > 0$. Given $T < \infty$, there exists a unique symmetric equilibrium, characterized by $\tilde{T} \in [0, T)$, in which the level of effort is given by*

$$u^*_{i,t} = \begin{cases} \frac{r(\alpha^{-1}-1)}{1+\frac{(1-\bar{p})\alpha}{\bar{p}-\alpha}e^{r(\alpha^{-1}-1)t}} & for\ t < \tilde{T} \\ 1 & for\ t \in [\tilde{T}, T] \end{cases}$$

The time $\tilde{T}$ is nondecreasing in the parameter T and strictly increasing for T large enough. Moreover, the posterior belief at time T strictly exceeds α.

See Figure 3-1 (provided in their paper) for illustration. Effort is first decreasing over time when the deadline is far enough in the future not to affect the agents' incentives (in

[3]For the purpose of this paper we just focus on this specific linear form of production technology. Bonatti and Horner (2011) also studies alternative form instantaneous probability to capture different model assumptions (e.g. synergies of cooperation).

[4]See Lemma 2 in Bonatti and Horner (2011).

(Parameter Values: $T = 3$, $r = 0.1$, $\alpha = 0.25$, $\bar{p} = 0.75$)

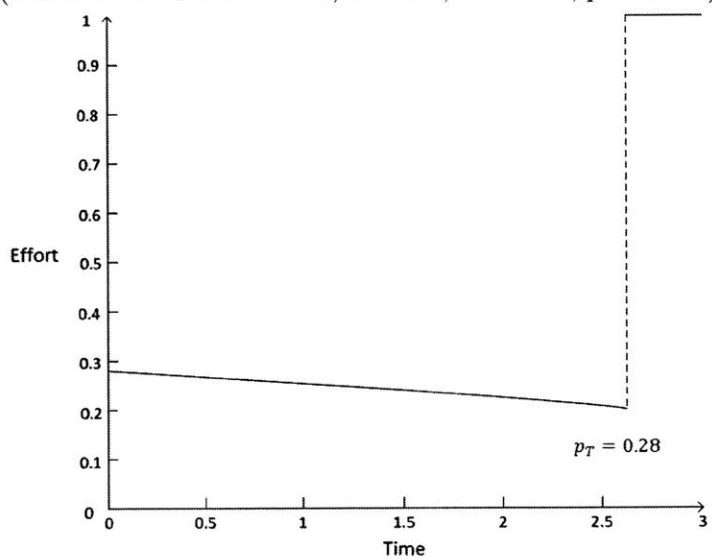


Figure 3-1. Optimal Strategies Given A Deadline of T=3

fact at this stage they behave as if there is no deadline). At some point $\hat{T}$ it is as if agents suddenly recognize that the deadline is near and exert maximal effort from then on. When the deadline comes if there is still no breakthrough, the posterior is still higher than α and agents would like to extend the deadline ex post.

The assumption regarding the discount rate (i.e. $\frac{1}{r} \geq \alpha^{-1} - \bar{p}^{-1}$) is to ensure that the agent's individual effort characterized in Theorem 1 is less than the maximum effort level 1. If this assumption is not satisfied, both agents exert maximal effort initially until at some point t_0 when the posterior belief p_{t_0} satisfies $r(\alpha^{-1} - p_{t_0}^{-1}) = 1$. From then on Theorem 1 applies. [5]

3.2.2 From Continuous-Time to Discrete-Time

A two-period model is introduced in Bonatti and Horner (2011) as "a simple example" to illustrate some of the key ideas in their paper. Their version of two-period setup can be summarized as follows:

Two agents collaborate on a project for two periods $T = 1, 2$. In each period agent i

[5]This discussion about the violation of the assumption appears in an earlier version of their paper, i.e. Horner and Bonatti (2009).

$(i = 1, 2)$ chooses effort level $u_{i,t} \in [0, \lambda]$; conditional on a good state, breakthrough happens in period t with probability $\Sigma_i u_{i,t}$. Individual cost of effort is $c(u_{i,t}) = \alpha u_{i,t}$.[6] The common prior that the project is good is $\bar{p}$ and agents have a common discount rate $\delta < 1$.

Bonatti and Horner (2011) set up the two-period model without solving for equilibrium. However, as we will see in Section 3.1, equilibrium results for this two-period game are not quite analagous to Theorem 1. This result is surprising, because one usually expects discrete-time models to be approximations for the continuous-time model.

Before delving into the solution details, we first ask the question of how to properly transform a continuous-time model into a discrete one. Consider the continuous-time model as what happens in the "real world", but we want to work with discrete-time models. What we can do is to divide the whole duration of the game equally into several periods, and set up a discrete-time game that roughly "approximate" what happens in continuous-time: specifically, we "tie up" the hands of the agents within each period by asking them to choose a constant instantaneous effort level throughout that period. Suppose the length of each period is chosen to be Δ. Each agent's choice variable in each discrete period is some level of instantaneous effort level $u_{i,t} \in [0, 1]$ that must stay unchanged during the period, and the corresponding cost of effort is $\alpha u_{i,t}\Delta$ as described in the continuous-time model. The discount factor δ is naturally derived as $\delta = e^{-r\Delta}$.

The probability of success "arriving" within this period is, in the "real world", $1 - e^{1-\Sigma_i u_{i,t}\Delta}$. However, one may find this expression not convenient to work with. When Δ is very small, this probability of success can be approximated by, say, $\Sigma_i u_{i,t}\Delta$. This is exactly what Bonatti and Horner do in their two-period model. But this is not the only way approximate the "real" probability: $1 - \Pi_i(1 - u_{i,t}\Delta) = \Sigma_i u_{i,t}\Delta - (\Pi_i u_{i,t})\Delta^2$ seems like an equally plausible choice.

Once we settle on the assumption about success probability within a period, we can then transform the continuous-time model to a discrete-time one with its setup summarized as follows:

- Two agents collaborate on a project for $T^D = \frac{T}{\Delta}$ periods (choose Δ wisely so that T^D

[6] In their discussion they allow a general convex cost function; I restrict to linear cost since cost is linear in the continuous-time model.

is an integer).[7]

- In each period $t = 1, ..., T^D$, agent i chooses an effort level $u_{i,t} \in [0, \lambda]$, where $\lambda = \Delta$. (Note that here, with some abuse of notations, $u_{i,t}$ is no longer the instantaneous effort level in a continuous-time sense, but is the aggregate effort level for the agent in that period.) The cost of effort $u_{i,t}$ is $c_i(u_{i,t}) = \alpha u_{i,t}$. Efforts are unobservable to other agents.

- At each period, breakthrough happens with probability $f(u_t)$ (where $u_t = (u_{i,t})_i$) if the state is good. (Possible choices of $f(u_t)$ includes $f(u_t) = \Sigma_i u_{i,t}$ or $f(u_t) = 1 - \Pi_i(1 - u_{i,t})$.) Breakthrough is impossible if the state is bad. A breakthrough is worth 1 to each agent at that period. The project ends after a breakthrough. Conditional on the state being good, breakthroughs across periods are independent.

- Common prior $\bar{p} \leq 1$ that $\omega = G$. Assume $\alpha < \bar{p}$.

- Common discount rate $\delta = e^{-r\Delta} < 1$.

A discrete-time model generated in this way can thus be defined as a discrete-time version of the continuous-time game.

As we shall see later, different choices regarding the approximation of success probability in the discrete-time model would lead to different equilibrium behavior. I give two plausible choices above: $f(u_t) = \Sigma_i u_{i,t}$, and $f(u_t) = 1 - \Pi_i(1 - u_{i,t})$. For future reference I name them respectively as "additive technology" and "product technology":

Additive technology refers to the assumption $f(u) = u_1 + u_2$, i.e. within a period the two agents' efforts contribute to success probability additively. This technology has two features: Firstly, within a period, the other agent's effort will not affect one's marginal productivity. As we will see later, this feature makes intermediate effort in equilibrium only a "knife-edge" case for discrete-time models. Secondly, combined with the assumption that breakthroughs accross periods are independent, this technology makes "working together" attractive. To see this point, suppose $u_1 = u_2 = \frac{1}{2}$, and the state is known to be good. Then if the two agents

[7]The superscript in T^D stands for "dicrete-time". In later parts of the paper when the context is clear to be about discrete-time models, I suppress the superscript and slightly abuse the notation T to also denote the total number of periods in a discrete-time game.

work together in the same period, they can guarantee a success; if they work sequentially, there is probability of $\frac{1}{4}$ that there is no success.

Product technology refers to the assumption $f(u) = u_1+u_2-u_1u_2 = 1-(1-u_1)(1-u_2)$, i.e. within a period the two agents are still conducting independent experiments. Correspondingly this means two things: First, within a period, one's marginal productivity decreases with the other agent's effort. This gives rise to intermediate effort in the spirit of mixed strategies as we look at symmetric equilibrium. Secondly, the attractiveness of "working together" disappears.

3.3 Two-Period Models

In this section we analyze two-period models in detail. We consider both additive and product technology. We will see that the models have some equilibria that resemble the Bonatti-Horner equilibrium in some respects, but they also have equilibria that look different, and have different existence and multiplicity properties.

For a given belief about the other agent's strategy $\hat{u}_{-i} = (\hat{u}_{-i,1}, \hat{u}_{-i,2})$, agent i has his best response correspondence $u_i(\hat{u}_{-i})$. Solving for a symmetric equilibrium (u^*, u^*) is to find a strategy $u^* = (u_1^*, u_2^*)$ (where u_1^*, u_2^* represent effort levels in period 1 and 2 respectively) that satisfies $u^* \in u_i(u^*)$. Since we are mostly interested in symmetric equilibria, for the purpose of concise notation we use $u^* = (u_1^*, u_2^*)$ to represent the symmetric equilibrium (u^*, u^*). So for example, when talking about a symmetric equilibrium, $(0, \lambda)$ refers to one in which both agents exert zero effort in period 1 and λ effort in period 2.

3.3.1 Additive Technology

Given the belief $\hat{u}_{-i} = (\hat{u}_{-i,1}, \hat{u}_{-i,2})$, agent i finds his best response by solving the following problem:

$$max_{u_{i,1},u_{i,2}}\bar{p}\{u_{i,1}+\hat{u}_{-i,1}-\alpha u_{i,1}+\delta(1-u_{i,1}-\hat{u}_{-i,1})(u_{i,2}+\hat{u}_{-i,2}-\alpha u_{i,2})\}+(1-\bar{p})(-\alpha u_{i,1}-\delta\alpha u_{i,2})$$

We solve this problem with two stages. First, for a given $u_{i,1}$, we find $u^*_{i,2}(u_{i,1})$ that maximizes expected payoff. Second, find the optimal $u_{i,1}$. As the first stage, it is easy to see that

$$u^*_{i,2}(u_{i,1}) = \begin{cases} \lambda & if\ u_{i,1} < u^o_{i,1}(\hat{u}_{-i,1}) \\ 0 & if\ u_{i,1} > u^o_{i,1}(\hat{u}_{-i,1}) \\ [0,\lambda] & if\ u_{i,1} = u^o_{i,1}(\hat{u}_{-i,1}) \end{cases}$$

where $u^o_{i,1}$ solves

$$\rho(u^o_{i,1}, \hat{u}_{-i,1}) \equiv \frac{\bar{p}(1 - u^o_{i,1} - \hat{u}_{-i,1})}{1 - \bar{p} + \bar{p}(1 - u^o_{i,1} - \hat{u}_{-i,1})} = \alpha$$

That is, $u^o_{i,1}$ is the cutoff level of agent i's effort in $t = 1$ such that the consequent *posterior* $\rho(u^o_{i,1}, \hat{u}_{-i,1})$ is exactly α and the agent is indifferent about his $u_{i,2}$ when it comes to $t = 2$. If he worked harder than the cutoff level, he will be too pessimistic to work in the second period, and if he worked less than that in period 1, he will be optimistic enough to exert maximal effort in period 2. If $u^o_{i,1} < 0$, it means agent i is pessimistic enough with $\hat{u}_{-i,1}$ alone. If $u^o_{i,1} > \lambda$, it means even with maximal effort in period 1, agent i is still optimistic enough about the good state in period 2.

Now as the second stage we rewrite the problem as

$$max_{u_{i,1}} \bar{p}\{u_{i,1} + \hat{u}_{-i,1} - \alpha u_{i,1} + \delta(1 - u_{i,1} - \hat{u}_{-i,1})(u^*_{i,2}(u_{i,1}) + \hat{u}_{-i,2} - \alpha u^*_{i,2}(u_{i,1}))\}$$

$$+(1 - \bar{p})(-\alpha u_{i,1} - \delta\alpha u^*_{i,2}(u_{i,1})) \stackrel{def}{=} max_{u_{i,1}} h(u_{i,1})$$

Given the simple structure of this problem, it is easy to see that h is continuous and is differentiable when $u_{i,1} \neq u^o_{i,1}$:

$$\frac{\partial h}{\partial u_{i,1}} = \begin{cases} \bar{p} - \alpha - \delta\bar{p}(\lambda + \hat{u}_{-i,2} - \alpha\lambda) & if\ u_{i,1} < u^0_{i,1}(\hat{u}_{-i,1}) \\ \bar{p} - \alpha - \delta\bar{p}\hat{u}_{-i,2} & if\ u_{i,1} > u^0_{i,1}(\hat{u}_{-i,1}) \end{cases}$$

The interpretation of the above expressions is pretty straightforward. The marginal benefit of effort $u_{i,1}$ comes from increasing the probability of success in period 1 conditional on the good state. The marginal cost of effort $u_{i,1}$ comes from two sources: one is the direct cost of

effort α, and the other is that of decreasing the probability that the game enters the second stage conditional the good state, which will bring in expectation positive utility.

We focus on solving for symmetric pure-strategy equilibria. The following proposition shows the existence of a symmetric pure-strategy equilibrium by finding one for all possible parameter conditions:

Proposition 2. *There exists a symmetric pure-strategy equilibrium for the two-period game with additive technology.*

Proof. If $\frac{\bar{p}(1-2\lambda)}{1-\bar{p}+\bar{p}(1-2\lambda)} \geq \alpha$, this condition means both agents will surely exert maximal effort in the second period no matter how hard they work in the first period (for the knife-edge case $\frac{\bar{p}(1-2\lambda)}{1-\bar{p}+\bar{p}(1-2\lambda)} = \alpha$, agents are indifferent in period 2 conditional on maximal effort in period 1, and we just assume that they take maximal effort in period 2). Taking that as given, the agents solve for their optimal first-period effort, which is either 0 or λ (again ignoring the possible indifference in period 1 in some knife-edge cases). Thus in this scenario, either $(0, \lambda)$ or (λ, λ) is an equilibrium.

If $\frac{\bar{p}(1-2\lambda)}{1-\bar{p}+\bar{p}(1-2\lambda)} < \alpha$, then we claim that $(\lambda, 0)$ is a symmetric equilibrium. Suppose agent i believes that $\hat{u}_{-i} = (\lambda, 0)$. The condition ensures that $u^o_{i,1} < \lambda$. If $u^o_{i,1} < 0$, then $\frac{\partial h}{\partial u_{i,1}} = \bar{p} - \alpha - \delta\bar{p}\hat{u}_{-i,2} = \bar{p} - \alpha > 0$, thus agent i's best response is indeed $(\lambda, 0)$. If $0 \leq u^o_{i,1} < \lambda$, then agent i is essentially choosing between strategies $(\lambda, 0)$ and $(0, \lambda)$. The additive technology dictates that the former is better regardless of the magnitude of discounting because working together is beneficial. Formally, this is because

$$\bar{p}\{\lambda + \lambda - \alpha\lambda\} + (1 - \bar{p})(-\alpha\lambda) > \bar{p}\{\lambda + \delta(1 - \lambda)(\lambda - \alpha\lambda)\} + (1 - \bar{p})(-\delta\alpha\lambda)$$

Therefore $(\lambda, 0)$ is indeed a symmetric equilibrium. □

Comparing to Bonatti and Horner (2011), one is perhaps most surprised by the proposed equilibrium of $(\lambda, 0)$, i.e. both agents exert maximal effort in the first period and none in the second. In their continuous-time model, "agents exert too little effort, and exert it too late. In the hope that the effort of others will suffice, they work less than they should early on, postponing their effort to a later time". But in the two-period model with additive

technology, why do we have this equilibrium where agents exert maximal effort in the first period and none in the second, thus do not procrastinate at all?

Intuitively, this is because with additive technology, working together in the same period is particularly attractive compared to working separately across different period. By assuming that the probability of success in a period equals the sum of efforts in that period, we are assuming something strictly better than giving two chances to conduct independent experiments and hoping for at least one success; on the other hand, across periods agents are conducting exactly independent experiments. When considering the defer the effort to the second period, it turns out that the benefit of potentially saving effort cost in case success happens in the first period is not worth the cost of decreasing overall success probability. So given that the other agent works hard in the first period, I would prefer to work hard in the first period as well, thus have no incentive to procrastinate.

Also, contrary to the uniqueness result in Bonatti and Horner (2011), we no longer have uniqueness in this additive-technology version of two-period model. To illustrate, we next show some non-trivial equilibrium multiplicity for an open set of parameters:

Proposition 3. *In the two-period model with additive technology, if the parameters satisfy the following conditions:*

$\frac{\bar{p}(1-\lambda)}{1-\bar{p}\lambda} < \alpha;$

$\bar{p} > \alpha;$

$\bar{p} - \delta\bar{p}\lambda < \alpha,$

then both $(0, \lambda)$ *and* $(\lambda, 0)$ *are symmetric equilibria for the game.*

Proof. Condition 1 says that if agent i believes that the other agent works extremely hard in the first period (i.e. $\hat{u}_{-i,1} = \lambda$), he is already pessimistic enough not to exert any effort in the second period. Condition 2, combined with Condition 1, ensures that if agent i believes that $(\hat{u}_{-i,1}, \hat{u}_{-i,2}) = (\lambda, 0)$, he will exert maximal effort in the first period, because he believes nobody will work in period 2 but it is worthwhile to work in period 1. Thus the first two conditions ensure that $(\lambda, 0)$ is an equilibrium.

Condition 3 says that if agent i believes that $\hat{u}_{-i,2} = \lambda$, then $\frac{\partial h}{\partial u_{i,1}} < 0$ for all $u_{i,1} \neq u^o_{i,1}$, i.e. if agent i believes that the other agent works extremely hard in the second period,

he would like to exert no effort at all in the first period in order to free-ride the second-period payoff. This, combined with Condition 2, ensures that if he furthermore believes that $(\hat{u}_{-i,1}, \hat{u}_{-i,2}) = (0, \lambda)$, then agent i will indeed choose $(0, \lambda)$ as the best response. Thus $(0, \lambda)$ is an equilibrium. □

Proposition 2 is not meant to give a tight result (i.e. the conditions given are sufficient but not necessary for the above two equilibria to co-exist), but is only to illustrate the point that equilibrium multiplicity is a non-trivial result in this version of two-period model, in contrast to the uniqueness result in Bonatti and Horner (2011). Intuitively, the coexistence of the two equilibria is not surprising once we understsand the attractiveness to work together under additive technology - it is kind of like a coordination game with multiple equilibria.

3.3.2 Product Technology

With product technology, the agent's best-response problem is slightly different:

$$max_{u_{i,1},u_{i,2}}\bar{p}\{u_{i,1}+\hat{u}_{-i,1}-\hat{u}_{-i,1}u_{i,1}-\alpha u_{i,1}+\delta(1-u_{i,1}-\hat{u}_{-i,1}+\hat{u}_{-i,1}u_{i,1})(u_{i,2}+\hat{u}_{-i,2}-\hat{u}_{-i,2}u_{i,2}-\alpha u_{i,2})\}$$

$$+(1-\bar{p})\{-\alpha u_{i,1} - \delta\alpha u_{i,2}\}$$

Take $\hat{u}_{-i,1}$ and $\hat{u}_{-i,2}$ as given, and let $u^*_{i,2}(u_{i,1})$ be the "best continuation strategy" given $u_{i,1}$. Note that with product technology, the other agent's same-period effort affects one's marginal productivity. We thus can see that

$$u^*_{i,2}(u_{i,1}) = \begin{cases} \lambda & if\ u_{i,1} < u^o_{i,1}(\hat{u}_{-i,1}, \hat{u}_{-i,2}) \\ 0 & if\ u_{i,1} > u^o_{i,1}(\hat{u}_{-i,1}, \hat{u}_{-i,2}) \\ [0, \lambda] & if\ u_{i,1} = u^o_{i,1}(\hat{u}_{-i,1}, \hat{u}_{-i,2}) \end{cases}$$

where cutoff $u^o_{i,1}$ solves

$$\frac{\bar{p}(1-u^o_{i,1})(1-\hat{u}_{i,1})}{1-\bar{p}+\bar{p}(1-u^o_{i,1})(1-\hat{u}_{i,1})}(1-\hat{u}_{-i,2}) = \alpha$$

Similarly rewriting the problem as

$$max_{u_{i,1},u_{i,2}}\bar{p}\{u_{i,1}+\hat{u}_{-i,1}-\hat{u}_{-i,1}u_{i,1}-\alpha u_{i,1}+\delta(1-u_{i,1}-\hat{u}_{-i,1}+\hat{u}_{-i,1}u_{i,1})$$

$$(u^*_{i,2}(u_{i,1})+\hat{u}_{-i,2}-\hat{u}_{-i,2}u^*_{i,2}(u_{i,1})-\alpha u^*_{i,2}(u_{i,1}))\}$$

$$+(1-\bar{p})\{-\alpha u_{i,1}-\delta\alpha u^*_{i,2}(u_{i,1}))\} \stackrel{def}{=} h(u_{i,1})$$

We have

$$\frac{\partial h}{\partial u_{i,1}}=\begin{cases}\bar{p}(1-\hat{u}_{-i,1})-\alpha-\delta\bar{p}(1-\hat{u}_{-i,1})(\lambda+\hat{u}_{-i,2}-\hat{u}_{-i,2}\lambda-\alpha\lambda) & if\ u_{i,1}<u^0_{i,1}(\hat{u}_{-i,1},\hat{u}_{-i,2})\\ \bar{p}(1-\hat{u}_{-i,1})-\alpha-\delta\bar{p}(1-\hat{u}_{-i,1})\hat{u}_{-i,2} & if\ u_{i,1}>u^0_{i,1}(\hat{u}_{-i,1},\hat{u}_{-i,2})\end{cases}$$

Again we are interested in symmetric equilibria. As it turns out, with product technology we indeed have uniqueness of equilibrium.

The way we prove uniqueness is to exhaust all possible scenarios of equilibra, find their corresponding necessary and sufficient conditions on parameters, and show that any two pairs of necessary and sufficient condtions cannot overlap.

Lemma 16. *If (x_1, x_2) is an equilibrium of the two-period game with product technology, then:*

1. $x_1 \neq 0$;
2. *If $x_1 \in (0,\lambda)$, then $u^o_{i,1}(x_1,x_2) \geq \lambda$, and $x_2 = \lambda$.*

Therefore the game with product technology can only have four possible kinds of equilibria:(x,λ), (λ,λ), (λ,m),$(\lambda,0)$ (where $0<x,\ m<\lambda$).

Proof. See Appendix. □

Lemma 17. *In a two-period model with product technology, we have the following sets of necessary and sufficient conditions for each possible equilibrium:*

1. *(x,λ) (where $x \in (0,\lambda)$) is an equilibrium iff* $\begin{cases}\frac{\bar{p}(1-x)(1-\lambda)^2}{1-\bar{p}+\bar{p}(1-x)(1-\lambda)} \geq \alpha & ...(1)\\ \bar{p}(1-x)\{1-\delta[(2-\alpha)\lambda-\lambda^2]\}=\alpha & ...(2)\end{cases}$

2. (λ, λ) *is an equilibrium iff* $\begin{cases} \frac{\bar{p}(1-\lambda)^2}{1-\bar{p}+\bar{p}(1-\lambda)^2}(1-\lambda) \geq \alpha & ...(3) \\ \bar{p}(1-\lambda)\{1-\delta[(2-\alpha)\lambda-\lambda^2]\} \geq \alpha & ...(4) \end{cases}$

3. (λ, m) *(where* $m \in (0, \lambda)$*) is an equilibrium iff* $\begin{cases} \frac{\bar{p}(1-\lambda)^2(1-m)}{1-\bar{p}+\bar{p}(1-\lambda)^2} = \alpha & ...(5) \\ \bar{p}(1-\lambda)\{1-\delta[\lambda+m-m\lambda-\alpha\lambda]\} \geq \alpha & ...(6) \end{cases}$

4. $(\lambda, 0)$ *is an equilibrium iff* $\begin{cases} \frac{\bar{p}(1-\lambda)^2}{1-\bar{p}+\bar{p}(1-\lambda)^2} \leq \alpha & ...(7) \\ \bar{p}(2\lambda-\lambda^2)-\alpha\lambda \geq \bar{p}\lambda+\delta(1-\bar{p}\lambda)(\frac{\bar{p}(1-\lambda)}{1-\bar{p}\lambda}\lambda-\alpha\lambda) & ...(8) \end{cases}$

Proof. See Appendix. □

In the above lemma, for each kind of symmetric equilibria, the first condition in the corresponding pair of necessary and sufficient conditions is to ensure that an agent is indeed willing to work as described in the second period, and the second condition is to ensure that an agent will work as described in the first period.

Notice that here the $(\lambda, 0)$ equilibrium exists only when δ is sufficiently small. If agents are patient enough, then this cannot be an equilibrium, which is in contrast to the additive technology where $(\lambda, 0)$ can be an equilibrium even without discounting.

Proposition 4. *In a two-period model with product technology, if there exists an equilibrium, then the equilibrium is unique.*

Proof. We only need to show that any two of the four sets of conditions in Lemma 2 cannot co-exist.

Note that condition (3), (5) and (7) are mutually exclusive. Also, obviously (2) and (4) contradicts with each other. Furthermore,

$$\frac{\bar{p}(1-x)(1-\lambda)^2}{1-\bar{p}+\bar{p}(1-x)(1-\lambda)} = \frac{\bar{p}(1-\lambda)^2}{\frac{1-\bar{p}}{1-x}+\bar{p}(1-\lambda)} < \frac{\bar{p}(1-\lambda)^2}{1-\bar{p}+\bar{p}(1-\lambda)^2}$$

so (1) and (7) contradicts each other.

It only remains to show that (x, λ) and (λ, m) cannot co-exist. Suppose by contradiction that they do, so conditions (1), (2), (5) and (6) all holds for some set of parameters. (1) and

(5) would imply that $x < m$. From (2) and (6) we have

$$\bar{p}(1-x)\{1-\delta[(2-\alpha)\lambda-\lambda^2]\} = \alpha \leq \bar{p}(1-\lambda)\{1-\delta[\lambda+m-m\lambda-\alpha\lambda]\}$$

Since $x < m$ and $\lambda < 1$ we can further write

$$\bar{p}(1-x)\{1-\delta[(2-\alpha)\lambda-\lambda^2]\} < \bar{p}(1-\lambda)\{1-\delta[\lambda+x-x\lambda-\alpha\lambda]\}$$

Both sides of the above inequility are linear in δ and obviously it does not hold when $\delta = 0$. So it must hold for $\delta = 1$, which gives

$$(1-x)\{1-[(2-\alpha)\lambda-\lambda^2]\} \leq (1-\lambda)\{1-[\lambda+x-x\lambda-\alpha\lambda]\}$$

Rearranging the inequality will eventually give us $\alpha\lambda^2 < \alpha\lambda x$, which implies $\lambda < x$, a contradiction. □

There remains a question of whether an equilibrium exist. Now that we have all these NS (necessary and sufficient) condtions for all possible equilibria (see Lemma 2), the question of equilibrium existence is equivalent to one asking whether it is true that any set of parameters falls into one of those sets of necessary and sufficient conditions. Surprisingly, as it turns out, the answer is no. Proposition 4 below gives a set of sufficient (but not necessary) condtions for non-existence of symmetric equilibrium.

Proposition 5. *In the two-period model with product technology, if parameters satisfy the following conditions (which gives a non-empty set that contains open sets in the parameter space):*

$$\begin{cases} \frac{\bar{p}(1-\lambda)^2}{1-\bar{p}+\bar{p}(1-\lambda)^2} \leq \alpha & \text{...(7)} \\ \bar{p}(2\lambda-\lambda^2)-\alpha\lambda \leq \bar{p}\lambda+\delta(1-\bar{p}\lambda)(\frac{\bar{p}(1-\lambda)}{1-\bar{p}\lambda}\lambda-\alpha\lambda) & \text{...(9)} \end{cases}$$

then there does not exist a pure-strategy symmetric equilibrium.

Proof. Note that the first condition is just condition (7) in Lemma 2 and the second condition (call it condition (9)) is the negation of condition (8). One can see that condition (9) holds

as long as δ is close enough to 1, and (7) would hold if $\bar{p}$ is very close to α.

Since (8) does not hold, $(\lambda, 0)$ is not an equilibrium. Since (7) holds and (7) contradicts with (3), (5) and (1), none of the three other candidates could be equilibrium. Thus pure-strategy symmetric equilibrium does not exist. □

3.3.3 Comparison with the Continous-Time Model

Both of the above two models are meant to tell a discrete-time (two-period) story the continuous-time model. As it turns out, neither of them gives a completely analogous result. We now put together the results and do some comparisons.

Existence and Uniqueness of (Pure-strategy, Symmetric) Equilibrium. The continuous-time game has a unique symmetric equilibrium (Theorem 1). In the two-period game with *additive* technology, equilibrium exists (Proposition 1) but is not unique in many cases (Proposition 2). In the two-period game with *product* technology, the equilibrium may not exist in many cases that are not knife-edge (Proposition 4), but once there is an equilibrium, it is unique (Proposition 3).

Effort Pattern. In the continuous-time game with deadline, the unique symmetric equilibrium features the following effort pattern: At first effort decreases over time, then at some point when it comes close to the deadline, effort jumps to its maximal level. (Depending on the length of the horizon, the equilibrium may not have the first stage with decreasing effort.)

In the two-period game with *additive* technology, intermediate level of effort appears in equilibrium only in some knife-edge parameter cases. The most common candidates of equilibria are (λ, λ), $(\lambda, 0)$ and $(0, \lambda)$. ($(0, 0)$ is obviously exluded as long as $\bar{p} > \alpha$.) With the Bonatti-Horner results in mind, $(\lambda, 0)$ seems most surprising, since it no longer features the deadline effect on effort. It seems to me that this equilibrium arise mainly as a result of the additive technology assumption (i.e. "working together" is attractive). The $(0, \lambda)$ equilibrium is more similar to Bonatti-Horner's result but is not quite the same, in the sense that in a continuous-time equilibrium effort level never reaches zero.

In the two-period game with *product* technology, intermediate level of effort arises sort of in the spirit of mixed strategies and is no longer "knife-edge". There are four candidates for equilibrium: (x, λ), (λ, λ), (λ, m) and $(\lambda, 0)$. The first two can be viewed as analogous to the Bonatti-Horner results; the last two, (λ, m) and $(\lambda, 0)$, are not quite analogous again.

The Final Posterior Belief. In the continuous-time game, agents are very careful not to let the posterior belief fall below α at any point of time. In two-period games, however, sometimes in equilibrium agents work hard in the first period and end up with a posterior which is pessimistic enough to prevent any further effort in the second period.

3.4 Multiple-Period Models without Learning

Since the two-period models give equilibrium predictions sometimes quite different from the continuous-time model, we naturally want to know if the equilibrium predictions would become closer to the prediction in the continuous-time model once we let give the agents more opportunities to adjust their effort level. However, since the analysis of multi-period model with uncertainty about the state is likely to be quite complicated, here we only look at models without learning - that is, the state is known to be good.

Before looking at discrete-time models let us first see what is the continuous-time equilibrium when there is no learning. This can be easily done by applying Theorem 1. We write it down as a separate theorem in this section.

Theorem 10. *Suppose the state is known to be good, and* $\frac{1}{r} \geq \alpha^{-1} - 1 > 0$. *Given* $T < \infty$, *there exists a unique symmetric equilibrium, characterized by* $\tilde{T} \in [0, T)$, *in which the level of effort is given by*

$$u^*_{i,t} = \begin{cases} r(\alpha^{-1} - 1) & for\ t < \tilde{T} \\ 1 & for\ t \in [\tilde{T}, T] \end{cases}$$

The time $\tilde{T}$ is nondecreasing in the parameter T and strictly increasing for T large enough. Moreover, the posterior belief at time T strictly exceeds α.

If the assumption $\frac{1}{r} \geq \alpha^{-1} - 1$ is violated, then the unique equilibrium features $u^*_{i,t} = 1$ for all $t \in [0, T]$.

In sum, if agents are patient enough, and if the time horizon is long enough, then the continuous-game has a unique symmetric equilibrium with the feature that effort is at first at some constant intermediate level and then jumps to the maximal level at some point.

Next we turn to analyzing the discrete-time models with additive and product technology respectively.

3.4.1 Additive Technology

Solution for A Fixed T-period Model

Since there is no learning, the posterior is always $\rho = 1$. Thus actions in previous periods do not affect the continuation game. Therefore, we can solve the game using backward induction.

Since $\alpha < \bar{p} = 1$, in the last period everyone works hard: $u_{i,T} = \lambda$. In eye of agent i , the value of entering period T is:

$$V_{i.T} = 2\lambda - \alpha\lambda$$

Now we look at period $T-1$, knowing what will happen in period T. Agent i's problem:

$$max_{u_{i,T-1}} \hat{u}_{-i,T-1} + u_{i,T-1} - \alpha u_{i,T-1} + \delta(1 - u_{i,T-1} - \hat{u}_{-i,T-1})V_{i,T}$$

The solution is straightforward:

If $\alpha + \delta V_{i,T} < 1$, then $u_{i,T-1} = \lambda$ and $V_{i,T-1} = 2\lambda - \alpha\lambda + \delta(1 - 2\lambda)V_{i,T}$;

If $\alpha + \delta V_{i,T} > 1$ then $u_{i,T-1} = 0$ and $V_{i,T-1} = \delta V_{i,T} < V_{i,T}$;

If $\alpha + \delta V_{i,T} = 1$, then backward induction does not leave a unique solution in period $T-1$ and we have multiple equilibria.

Similar analysis applies to all previous periods. In sum, the above analysis points directly to the following conclusion:

(Parameter Values: $T = 7$, $\delta = 0.9$, $\alpha = 0.4$, $\lambda = 0.3$)

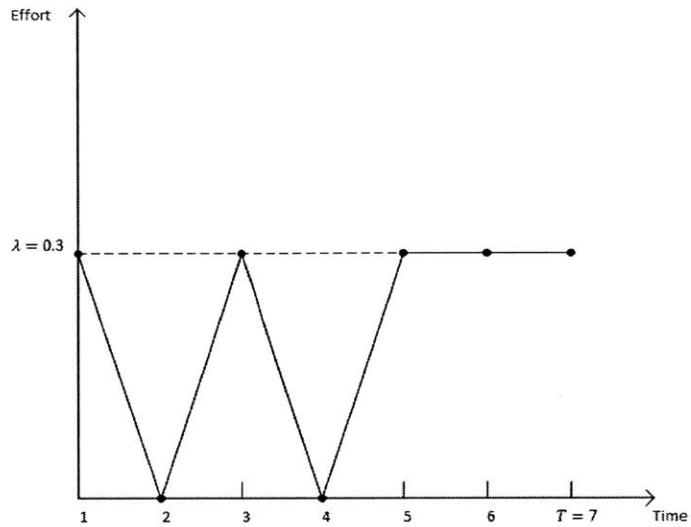


Figure 3-2. Equilibrium Individual Effort with Additive Technology

Proposition 6. *In a T-period game with additive technology, if $\bar{p} = 1$, an equilibrium exists, and is generically unique and symmetric, in which case the equilibrium effort level in each period is extremal (i.e. either λ or 0). If $\delta \leq \frac{1-\alpha}{1-\alpha+\alpha\lambda}$, then $V_{i,t} < \frac{1-\alpha}{\delta}$ for all t and the equilibrium effort is always maximal no matter how large T is: $u^*_{i,t} = \lambda$ for all t. If $\delta > \frac{1-\alpha}{1-\alpha+\alpha\lambda}$, and if we are in a generic case where the equilibrium is unique, then there exists some period $\hat{T} < T$ such that in equilibrium, $u^*_{i,t} = \lambda$ for all $t \geq \hat{T}$, and for $t < \hat{T}$, the effort level is alternating between 0 and λ, i.e. one (or several) period(s) with zero effort followed and preceded by one (or several) period(s) of maximal effort.*

Proof. See Appendix. □

Figure 3-2 below shows the equilibrium effort path we solve for a game with the following set of parameters: $\delta = 0.9$, $T = 7$, $\alpha = 0.4$, $\lambda = 0.3$. After period 5, the effort level stays at λ. Before that, the effort level alternates between λ and 0.

It is worth pointing out that the condition $\delta > \frac{1-\alpha}{1-\alpha+\alpha\lambda}$ is a discrete-time version of the assumption $\frac{1}{r} \geq \alpha^{-1} - 1$. This can be checked by recalling the relationship $\delta = e^{-r\Delta}$ and $\lambda = \Delta$, and then let Δ goes to zero.

Without learning, the equilibrium in discrete-time model with additive technology is already quite similar to the continuous-time case: the symmetric equilibrium exists and is

generically unique, and we see maximal effort when it comes close to the deadline, which is a feature of the Bonatti-Horner model with deadline.

However, one would quickly observe the difference: In a discrete-time model without learning, equilibrium effort level is generically extremal, and in those non-generic cases where intermediate level of effort may appear, there are certainly multiple equilibria. In contrast, in continuous-time model, symmetric equilibrium is unique, and during the early stage the equilibrium effort is at some intermediate level.

Also, if we look at asymmetric equilibria, we shall note that generically there is no asymmetric equilibrium in a discrete-time model with additive technology, but Bonatti and Horner (2011) show that there are asymmetric equilibria in the continuous-time game.

It seems that one "special" equilibrium in a non-generic case of the dicrete-time model with additive technology gives the best analogy of the continuous-time equilibrium. Suppose the parameters are such that $2\lambda - \alpha\lambda = V_{i,T} < V_{i,T-1} < ... < V_{i,\hat{T}} = \frac{1-\alpha}{\delta}$, i.e. backward induction gives unique solution up till period $\hat{T}$, but at $\hat{T}-1$, since the value of entering period $\hat{T}$ is exactly $\frac{1-\alpha}{\delta}$, agents are indifferent among all their effort levels. In a symmetric equilibrium, suppose $u_{i,\hat{T}-1} = x$, then $V_{i,\hat{T}-1} = (2-\alpha)x + \delta(1-2x)V_{i,\hat{T}}$. There is one particular effort level that keeps $V_{i,\hat{T}-1}$equal to $V_{i,\hat{T}}$: It solves: $(2-\alpha)x+\delta(1-2x)\frac{1-\alpha}{\delta} = \frac{1-\alpha}{\delta}$. This effort level is $x = \frac{1-\delta}{\delta}\frac{1-\alpha}{\alpha}$. Thus $u_{i,t} = \frac{1-\delta}{\delta}\frac{1-\alpha}{\alpha}$ for all $t < \hat{T}$ and $u_{i,t} = \lambda$ for all $t \geq \hat{T}$ constitutes an equilibrium, hereinafter referred to as the "special" equilibrium.

The above "special" equilibrium looks exactly like the equilibrium in the continuous-time model: constant intermediate level of effort at some first stage and then maximal effort in the second stage. However such equilibrium does not exist for a generic set of parameters for the discrete-time model. The reason is that, in order to have intermediate level of effort in equilibrium, the agent must be indifferent about his effort in that period, which requires the continuation value of the game to be exactly $V_{i,t} = \frac{1-\alpha}{\delta}$; however in a discrete-time model, generically the coincidence that $V_{i,\hat{T}} = \frac{1-\alpha}{\delta}$ does not happen and we usually end up "jumping over the indifference point". Therefore no matter how hard we try to approximate the continuous-time game (i.e. use a discrete-time model with very short periods), generically we cannot have the ideally analogous equilibrium.

To the contrary, one can imagine heuristically that in a continuous-time model, as we do

the backward induction[8], $V_{i,t}$ decreases (as time goes backwards from $t = T$) *continuously*, thus it will surely cross the value that makes agents indifferent about their effort levels in the previous instant. Thus in a continuous-time model we will surely have the "special" equilibrium, which is indeed the unique symmetric equilibrium there.

3.4.2 Product Technology

Now we look at models with product technology. As before, without learning, the game is solvable by backward induction. If agent i can pin down what will happen if the game enters period $t+1$. and suppose his continuation value into period $t+1$ is $V_{i,t+1}$, then his problem at period t is the following program:

$$max_{u_{i,t}} \hat{u}_{-i,t} + u_{i,t} - \hat{u}_{-i,t}u_{i,t} - \alpha u_{i,t} + \delta(1 - \hat{u}_{-i,t} - u_{i,t} + \hat{u}_{-i,t}u_{i,t})V_{i,t+1}$$

where $\hat{u}_{-i,t}$ is agent i's belief about the other agent's effort in period t. Obviously the solution is as follows:

If $(1 - \hat{u}_{-i,t})(1 - \delta V_{i,t+1}) > \alpha$, then $u_{i,t} = \lambda$;

If $(1 - \hat{u}_{-i,t})(1 - \delta V_{i,t+1}) < \alpha$, then $u_{i,t} = 0$;

If $(1 - \hat{u}_{-i,t})(1 - \delta V_{i,t+1}) = \alpha$, then agent i is indifferent about his effort level in period t.

One can quickly see the existence and uniqueness of symmetric equilibrium from the above analysis. The next proposition characterizes the symmetric equilibrium in the model with product technology:

Proposition 7. *In a T-period model with product technology, if $\bar{p} = 1$, there exists a unique symmetric equilibrium. If $\delta \leq \frac{1-\lambda-\alpha}{(1-\lambda)(1-\alpha)}$, then $u^*_{i,t} = \lambda$ for all t, which consistitutes the only equilibrium. If $\delta > \frac{1-\lambda-\alpha}{(1-\lambda)(1-\alpha)}$, then in the unique* symmetric *equilibrium, the strategy is as follows:*

$$u^*_{i,t} = \begin{cases} \lambda & if\ \hat{T} < t \leq T \\ \frac{(1-\delta)(1-\alpha)}{1-\delta+\delta\alpha} & if\ 1 \leq t < \hat{T} \\ x \in (\frac{(1-\delta)(1-\alpha)}{1-\delta+\delta\alpha}, \lambda] & if\ t = \hat{T} \end{cases}$$

[8]If there is learning, the continuous-time model cannot be solved by backward induction, and this is indeed the difficulty of further discussion.

(Parameter Values: $T = 7$, $\delta = 0.9$, $\alpha = 0.4$, $\lambda = 0.3$)

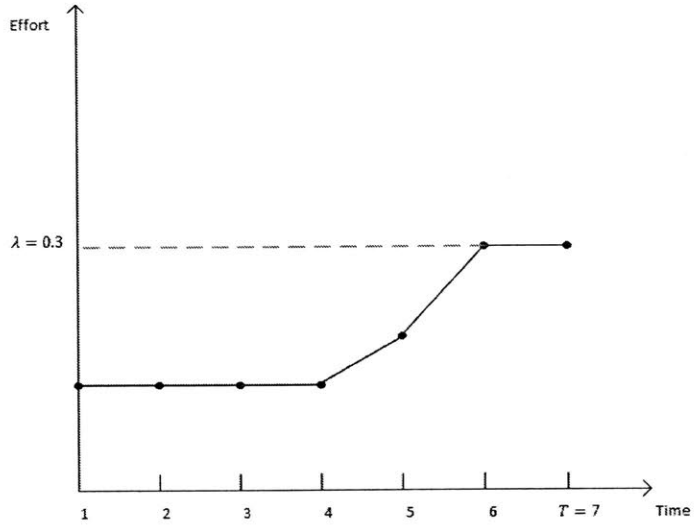


Figure 3-3. Equilibrium Individual Effort with Product Technology

Proof. See Appendix. □

Figure 3-3 shows the equilibrium effort path we solve for a game with the following set of parameters: $\delta = 0.9$, $T = 7$, $\alpha = 0.4$, $\lambda = 0.3$.

The equilibrium in this case is obviously very analogous to that in a continuous-time game. Furthermore, the condition $\delta > \frac{1-\lambda-\alpha}{(1-\lambda)(1-\alpha)}$ is again parallel to $\frac{1}{r} \geq \alpha^{-1} - 1$. There is even an easy way of checking "convergence":

$$\frac{(1-\delta)(1-\alpha)}{1-\delta+\delta\alpha} = \frac{(1-e^{-r\Delta})(1-\alpha)}{1-e^{-r\Delta}+e^{-r\Delta}\alpha} \to r\Delta \ (\Delta \to 0)$$

That is, as the period-length Δ goes to zero, the constant effort level in the first predicted by a discrete-time model converges to the level predicted by a continuous-time model. It is clearly a positive evidence that the limit of equilibrium in the discrete-time with product technology is indeed the equilibrium of the continuous-time game, despite the fact that we don't have a definition of "convergence" or "limit" of equilibrium.

3.5 Conclusion and Future Research

In this paper I examine several discrete-time versions of the continuous-time model of dynamic moral hazard in teams in Bonatti and Horner (2011). There are different ways to model production technology, both of which seems to approximate the technology in the continuous-time model but different choices of technology assumption may lead to quite different equilibrium prediction in discrete-time models.

The additive technology, perhaps unexpectedly, essentially depicts a kind of synergy of cooperation because under this assumption it is more attractive to work together in the same period than to work separately across different periods. Under the assumption of additive technology, in the two-period model we always have existence of symmetric equilibrium, but equilibrium multiplicity may arise for an open set of parameter values. In particular, regardless of the discounting factor, we may have a symmetric equilibrium in which agents exert maximal effort in the first period and none in the second, which seems quite different from the equilibrium prediction in the continuous-time model. In multi-period model without learning, generically we have a unique (symmetric) equilibrium in which agents will always exert extremal level of effort.

Under the assumption of product technology, in the two-period model we have uniqueness of symmetric equilibrium, but symmetric equilibrium may not exist under some parameter conditions, as opposed to the existence result in Bonatti and Horner (2011). In multi-period model without learning, we have a unique symmetric equilibrium that is quite analagous to the continuous-time model prediction.

The different versions of discrete-time models provide an alternative perspective in understanding the collaborating game. The analysis suggests that perhaps one should not assume that discrete and continuous-time models would always tell similar stories and give analogous predictions, especially in strategic settings. If we believe that the real world game is closer to a discrete-time setting, we may want to be cautious when interpreting the equilibrium results in a continuous-time model.

Finally, a theoretically intriguing question is how multi-period model with learning would behave, especially when the number of periods increases and the discrete-time setup approx-

imates the continuous-time model increasingly better. Along that way we may understand how equilibrium multiplicity in the two-period additive technology model and equilibrium non-existence in the two-period product technology model eventually turn into existence and uniqueness in the continuous-time model. That would be a direction for future research.

Appendix

Proof of Lemma 1 1. Firstly, obviously $(0,0)$ is not an equilibrium as long as $\bar{p} > \alpha$. We then show that $(0,x)$ (where $x > 0$) is also not an equilibrium: Given a belief of $(\hat{u}_{-i,1}, \hat{u}_{-i,2}) = (0,x)$, we show that $(u_{i,1}, u_{i,2}) = (x,0)$ yields higher payoff than $(u_{i,1}, u_{i,2}) = (0,x)$. If $(u_{i,1}, u_{i,2}) = (x,0)$, the payoff is

$$\bar{p}\{x - \alpha x + \delta(1-x)x\} + (1-\bar{p})\{-\alpha x\} = \bar{p}x - \alpha x + \delta\bar{p}(1-x)x$$

If $(u_{i,1}, u_{i,2}) = (0,x)$, the payoff is

$$\bar{p}\{\delta(2x - x^2 - \alpha x)\} + (1-\bar{p})\{-\delta\alpha x\} = \delta(\bar{p}x - \alpha x) + \delta\bar{p}(1-x)x$$

Obviously the former is higher if $\delta < 1$.

2. If (x_1, x_2) is an equilibrium and $x_1 \in (0,\lambda)$, first we show that $u^0_{i,1}(\hat{u}_{-i,1}, \hat{u}_{-i,2}) \geq \lambda$ (where $(\hat{u}_{-i,1}, \hat{u}_{-i,2}) = (x_1, x_2)$). Suppose first that $u^0_{i,1}(\hat{u}_{-i,1}, \hat{u}_{-i,2}) \in (0,\lambda)$, since

$$\bar{p}(1-\hat{u}_{-i,1}) - \alpha - \delta\bar{p}(1-\hat{u}_{-i,1})(\lambda + \hat{u}_{-i,2} - \hat{u}_{-i,2}\lambda - \alpha\lambda) < \bar{p}(1-\hat{u}_{-i,1}) - \alpha - \delta\bar{p}(1-\hat{u}_{-i,1})\hat{u}_{-i,2}$$

we can see that at the optimum u_1 should never take any intermediate level of effort. Contradiction.

If $u^0_{i,1}(\hat{u}_{-i,1}, \hat{u}_{-i,2}) \leq 0$, meaning given $\hat{u}_{-i,1}$ and $\hat{u}_{-i,2}$, agent i is always so pessimistic about the state that he doesn't want to exert any effort in period 2 regardless of $u_{i,1}$ (at least for intermediate $u_{i,1}$), then by symmetry $\hat{u}_{-i,2} = x_2 = 0$; and since $u_1 = x_1 \in (0,\lambda)$ we must have $\bar{p}(1-\hat{u}_{-i,1}) - \alpha = 0$. But now $\frac{\bar{p}(1-\hat{u}_{-i,1})}{1-\bar{p}\hat{u}_{-i,1}} > \bar{p}(1-\hat{u}_{-i,1}) = \alpha$, meaning that given $\hat{u}_{-i,1}$ and $\hat{u}_{-i,2} = 0$, agent i should be willing to work if $u_{i,1} = 0$, contradicting the assumption.

So the only possibility is $u^0_{i,1}(\hat{u}_{-i,1}, \hat{u}_{-i,2}) \geq \lambda$, meaning that given $\hat{u}_{-i,1}$ and $\hat{u}_{-i,2}$, agent i is always optimistic enough to exert effort λ regardless of his choice of $u_{i,1}$(at least for intermediate $u_{i,1}$). By symmetry $\hat{u}_{-i,2} = x_2 = \lambda$.

Proof of Lemma 2 1. If (x, λ) is an equilibrium, in the proof of Lemma 1 we show that $u^0_{i,1}(x, \lambda) \geq \lambda$, meaning that given the belief that $(\hat{u}_{-i,1}, \hat{u}_{-i,2}) = (x, \lambda)$, in the second period agent i is always optimistic enough to exert highest effort. This translates directly into condition (2). Furthermore, we need $\frac{\partial h}{\partial u_{i,1}} = 0$ so that $u_{i,1} = x \in (0, \lambda)$ is indeed an optimal choice for agent i, which translates into condition (1).

Conversely, if conditions (1) and (2) are satisfied, we show that (x, λ) is indeed an equilibrium. Given the belief $(\hat{u}_{-i,1}, \hat{u}_{-i,2}) = (x, \lambda)$, condition (2) $\Rightarrow u^0_{i,1}(x, \lambda) \geq \lambda$ and $u^*_{i,2}(u_{i,1}) = \lambda$ for all $u_{i,1} \in [0, \lambda)$. Condition (1) further ensures that $\frac{\partial h}{\partial u_{i,1}} = 0$ for all $u_{i,1} \in [0, \lambda)$, which proves that $u_i = (x, \lambda)$ is indeed a best response to the belief (x, λ).

2. If (λ, λ) is an equilibrium, then firstly, given the belief $(\hat{u}_{-i,1}, \hat{u}_{-i,2}) = (\lambda, \lambda)$, it must be that $\lambda \in u^*_{i,2}(\lambda)$, which translates into condition (3). Furthermore this says $u^*_{i,2}(u_{i,1}) = \lambda$ for all $u_{i,1} \in [0, \lambda)$, thus $\frac{\partial h}{\partial u_{i,1}} = \bar{p}(1-\lambda)\{1 - \delta[(2-\alpha)\lambda - \lambda^2]\} - \alpha$ for all $u_{i,1} \in [0, \lambda)$. Since $u_{i,1} = \lambda$ is an optimal choice, we need $\frac{\partial h}{\partial u_{i,1}} \geq 0$, which is condition (4).

Conversely, it is easy to see that (3) and (4) ensures that (λ, λ) is indeed an equilibrium.

3. Given the belief (λ, m), condition (5) says that $u^*_{i,2}(\lambda) = [0, \lambda]$, which implies that $u^*_{i,2}(u_{i,1}) = \lambda$ for all $u_{i,1} \in [0, \lambda)$. Condition (6) ensures that $u_{i,1} = \lambda$ is the optimal choice, given the belief (λ, m) and that $u^*_{i,2}(u_{i,1}) = \lambda$ for all $u_{i,1} \in [0, \lambda)$. It is then easy to see that (5) and (6) combined are necessary and sufficient for (λ, m) to be an equilibrium.

4. Given the belief $(\hat{u}_{-i,1}, \hat{u}_{-i,2}) = (\lambda, 0)$, condition (7) says that $0 \in u^*_{i,2}(\lambda)$, and condition (8) says that $u_i = (\lambda, 0)$ yields higher payoff to agent i than $u_i = (0, \lambda)$. Obviously (7) and (8) are necessary for $(\lambda, 0)$ to be an equilibrium. They are also sufficient: to show that $u_i = (\lambda, 0)$ is indeed the best response to $\hat{u}_{-i} = (\lambda, 0)$, we only need to show: A) $0 \in u^*_{i,2}(\lambda)$, which is equivalent to condition (7), and B) observe that $h(u_{i,1})$ is convex, meaning either $u_{i,1} = 0$ or $u_{i,1} = \lambda$ must be optimal, so we only need $u_i = (\lambda, 0)$ to be better than both $u_i = (0, \lambda)$ and $u_i = (0, 0)$. Condition (8) ensures that $(\lambda, 0)$ is better than $(0, \lambda)$, while $\bar{p} > \alpha$ guarantees that $(\lambda, 0)$ is better than $(0, 0)$.

Proof of Proposition 5 Recall the backward induction procedure:

1. $u_{i,T} = \lambda$, $V_{i,T} = 2\lambda - \alpha\lambda$.
2. For $t < T$:

- If $\alpha + \delta V_{i,t+1}<1$, then $u_{i,t} = \lambda$ and $V_{i,t} = 2\lambda - \alpha\lambda + \delta(1-2\lambda)V_{i,t+1}$;
- If $\alpha + \delta V_{i,t+1}>1$ then $u_{i,t} = 0$ and $V_{i,t} = \delta V_{i,t+1} < V_{i,t+1}$;
- If $\alpha + \delta V_{i,t+1}=1$, then backward induction does not pin down a unique solution for effort in period t and we have multiple symmetric equilibria. $u_{i,t} = x \in [0, \lambda]$ and $V_{i,t} = 2x - \alpha x + \delta(1-2x)V_{i,t+1}$ would work as one of the symmetric solutions.

Existence of symmetric equilibrium can be seen from the backward induction procedure: the game is symmetric, and backward induction always gives a solution at every step. Equilibrium multiplicity only arises when $\alpha + \delta V_t=1$ for some period t, which generically does not hold. The remaining claims are obvious given the backward induction procedure.

Note that $V_{i,t-1} = 2\lambda - \alpha\lambda + \delta(1-2\lambda)V_{i,t} \geq V_{i,t}$ iff $V_{i,t} \leq \frac{1}{1-\delta(1-2\lambda)}(2\lambda - \alpha\lambda)$. If $\delta \leq \frac{1-\alpha}{1-\alpha+\alpha\lambda}$, then one can verify that this is equivalent to $\frac{1}{1-\delta(1-2\lambda)}(2\lambda - \alpha\lambda) \leq \frac{1-\alpha}{\delta}$, so starting from $V_{i,T} = 2\lambda - \alpha\lambda < \frac{1}{1-\delta(1-2\lambda)}(2\lambda - \alpha\lambda)$, it is easy to show by induction that for any $t < T$ we have $V_{i,t+1} < V_{i,t} < \frac{1}{1-\delta(1-2\lambda)}(2\lambda - \alpha\lambda) \leq \frac{1-\alpha}{\delta}$, thus $u_{i,t} = \lambda$ for all t.

If $\delta > \frac{1-\alpha}{1-\alpha+\alpha\lambda}$, then $\frac{1}{1-\delta(1-2\lambda)}(2\lambda - \alpha\lambda) > \frac{1-\alpha}{\delta}$. Starting from $V_{i,T} = 2\lambda - \alpha\lambda$ it is easy to show that in a generic case (i.e. $V_{i,t} = \frac{1-\alpha}{\delta}$ never holds), there exists some $\hat{T} < T$ such that $V_{i,t} < \frac{1-\alpha}{\delta}$ for all $t > \hat{T}$ and $V_{i,\hat{T}} > \frac{1-\alpha}{\delta}$. Thus $u_{i,t} = \lambda$ for all $t \geq \hat{T}$ and $u_{i,\hat{T}-1} = 0$. Since the backward induction dictates that $V_{i,t-1} = \delta V_{i,t} < V_{i,t}$ if $V_{i,\hat{T}} > \frac{1-\alpha}{\delta}$, we know $V_{i,\hat{T}-1} = \delta V_{i,\hat{T}} < V_{i,\hat{T}}$, and we know this decreasing trend can only last for a limited number of periods and at some point (say $\tilde{T} < \hat{T}$) we must have $V_{i,\tilde{T}} < \frac{1-\alpha}{\delta}$ and we enter another increasing trend. Therefore before period $\hat{T}$ the effort level is alternating between 0 and λ in a generic case.

Proof of Proposition 6 To show existence and uniqueness of a symmetric equilibrium, we only need to show the existence and uniqueness of a symmetric solution in every step of backward induction. Indeed, at period t, if $(1-\lambda)(1-\delta V_{i,t+1}) > \alpha$, then $u_{i,t} = u_{-i,t} = \lambda$ is the only solution at period t given by backward induction; if $(1-\delta V_{i,t+1}) < \alpha$, then $u_{i,t} = u_{-i,t} = 0$ is the only solution; if $1 - \delta V_{i,t+1} > \alpha > (1-\lambda)(1-\delta V_{i,t+1})$, then $u_{i,t} = u_{-i,t} = 1 - \frac{\alpha}{1-\delta V_{i,t+1}}$ gives the only symmetric solution (in this case there are also asymmetric solution in this step, leading to asymmetric equilibria).

If $\delta \leq \frac{1-\lambda-\alpha}{(1-\lambda)(1-\alpha)}$, then starting from some $V_{i,t+1}$ given by backward induction, if $V_{i,t+1} < \frac{1-\frac{\alpha}{1-\lambda}}{\delta}$ (which holds when $t = T$), then we must have $u_{i,t} = u_{-i,t} = \lambda$ and $V_{i,t} = 2\lambda - \lambda^2 - \alpha\lambda + \delta(1-\lambda)^2 V_{i,t+1} < 2\lambda - \lambda^2 - \alpha\lambda + \delta(1-\lambda)^2 \frac{1-\frac{\alpha}{1-\lambda}}{\delta} = 1 - \alpha \leq \frac{1-\frac{\alpha}{1-\lambda}}{\delta}$. Thus by induction $V_{i,t+1} < \frac{1-\frac{\alpha}{1-\lambda}}{\delta}$ for all t thus $u^*_{i,t} = \lambda$ for all t is the only equilibrium.

If $\delta > \frac{1-\lambda-\alpha}{(1-\lambda)(1-\alpha)}$, then in a symmetric equilibrium there must exists some $\hat{T}$ s.t. $0 = V_{i,T+1} < V_{i,T} < ... < V_{i,\hat{T}+2} < \frac{1-\frac{\alpha}{1-\lambda}}{\delta}$ thus $u^*_{i,t} = \lambda$ for all $T > \hat{T}$, but $V_{i,\hat{T}+1} = 2\lambda - \lambda^2 - \alpha\lambda + \delta(1-\lambda)^2 V_{i,\hat{T}+2} \in [\frac{1-\frac{\alpha}{1-\lambda}}{\delta}, 1-\alpha)$. This means $u^*_{i,\hat{T}} = 1 - \frac{\alpha}{1-\delta V_{i,\hat{T}+1}} \in (\frac{(1-\delta)(1-\alpha)}{1-\delta+\delta\alpha}, \lambda]$ and furthermore $V_{i,\hat{T}} = 2u^*_{i,\hat{T}} - (u^*_{i,\hat{T}})^2 - \alpha u^*_{i,\hat{T}} + \delta(1-u^*_{i,\hat{T}})^2 V_{i,\hat{T}+1} = 1-\alpha$. This further indicates that $u^*_{i,t} = \frac{(1-\delta)(1-\alpha)}{1-\delta+\delta\alpha}$ and $V^*_{i,t} = 1-\alpha$ for all $t < \hat{T}$.

Bibliography

Afèche, Philipp, Zhe Liu, and Costis Maglaras. 2018. "Ride-hailing networks with strategic drivers: The impact of platform control capabilities on performance." *SSRN: https://ssrn.com/abstract=3120544* . 70

Banerjee, Siddhartha, Carlos Riquelme, and Ramesh Johari. 2015. "Pricing in ride-share platforms: A queueing-theoretic approach." *SSRN: https://ssrn.com/abstract=2568258* . 69

Bergemann, Dirk and Stephen Morris. 2005. "Robust Mechanism Design." *Econometrica* 73 (6):1771–1813. 16

Bertsekas, Dimitri P. 2005. *Dynamic programming and optimal control*, vol. 1. Athena scientific Belmont, MA. 77, 79, 80

Besbes, Omar and Assaf Zeevi. 2009. "Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms." *Operations Research* 57 (6):1407–1420. 70

Biais, Bruno, Thomas Mariotti, Guillaume Plantin, and Jean-Charles Rochet. 2007. "Dynamic security design: Convergence to continuous time and asset pricing implications." *The Review of Economic Studies* 74 (2):345–390. 138

Bonatti, Alessandro and Johannes Horner. 2011. "Collaborating." *The American Economic Review* 101 (2):632–663. 135, 136, 137, 138, 139, 140, 141, 145, 146, 147, 155, 158

Carroll, Gabriel. 2015. "Robustness and Linear Contracts." *American Economic Review* 105 (2):536–563. 14, 15, 17, 20, 40, 41, 42, 44, 46, 56

———. 2017. "Robustness and Separation in Multi-Dimensional Screening." *Econometrica* 85 (2):453–488. 16

Carroll, Gabriel and Ilya Segal. 2017. "Robustly Optimal Auctions with Unknown Resale Opportunities." Tech. rep., Stanford University. 16

Carvalho, Alexandre X and Martin L Puterman. 2005. "Learning and pricing in an internet environment with binomial demands." *Journal of Revenue and Pricing Management* 3 (4):320–336. 70

Castillo, Juan Camilo, Dan Knoepfle, and Glen Weyl. 2017. "Surge pricing solves the wild goose chase." In *Proceedings of the 2017 ACM Conference on Economics and Computation*. ACM, 241–242. 69

Chassang, Sylvain. 2013. "Calibrated Incentive Contracts." *Econometrica* 81 (5):1935–1971. 16, 42

Che, Yeon-Koo and Seung-Weon Yoo. 2001. "Optimal Incentives for Teams." *American Economic Review* 91 (3):525–541. 16

Chung, Kim-Sau and J. Ely. 2007. "Foundations of Dominant-Strategy Mechanisms." *Review of Economic Studies* 74 (2):447–476. 16

DeMarzo, Peter M and Yuliy Sannikov. 2006. "Optimal security design and dynamic capital structure in a continuous-time agency model." *The Journal of Finance* 61 (6):2681–2724. 138

Figliozzi, Miguel Andres, Hani S Mahmassani, and Patrick Jaillet. 2007. "Pricing in dynamic vehicle routing problems." *Transportation Science* 41 (3):302–318. 69

Frankel, Alexander. 2014. "Aligned Delegation." *American Economic Review* 104 (1):66–83. 16

Gallego, Guillermo. 1992. "A minmax distribution free procedure for the (Q, R) inventory model." *Operations Research Letters* 11 (1):55–60. 92, 132

Gallego, Guillermo and Garrett Van Ryzin. 1994. "Optimal dynamic pricing of inventories with stochastic demand over finite horizons." *Management science* 40 (8):999–1020. 70

———. 1997. "A multiproduct dynamic pricing problem and its applications to network yield management." *Operations research* 45 (1):24–41. 70

Garicano, Luis and Tano Santos. 2004. "Referrals." *American Economic Review* 94 (3):499–525. 16, 47

Garrett, Daniel. 2014. "Robustness of Simple Menus of Contracts in Cost-based Procurement." *Games and Economic Behavior* 87:631–641. 16

Georgiadis, George. 2014. "Projects and team dynamics." *The Review of Economic Studies* 82 (1):187–218. 137

Guo, Yingni. 2016. "Dynamic delegation of experimentation." *American Economic Review* 106 (8):1969–2008. 137

Hackman, Richard. 2002. *Leading Teams: Setting the Stage for Great Performances*. Boston, MA: Harvard Business School Press. 16, 19

Halac, Marina, Navin Kartik, and Qingmin Liu. 2016. "Optimal contracts for experimentation." *The Review of Economic Studies* 83 (3):1040–1091. 137

Holmström, Bengt. 1982. "Moral Hazard in Teams." *The Bell Journal of Economics* 13 (2):324–340. 13, 14, 20

Holmström, Bengt and Paul Milgrom. 1987. "Aggregation and Linearity." *Econometrica* 55 (2):303–328. 15

Horner, Johannes and Alessandro Bonatti. 2009. "Collaborating." In *2009 Meeting Papers*, 1019. Society for Economic Dynamics. 140

Hurwicz, Leonid and Leonard Shapiro. 1978. "Incentive Structures Maximizing Residual Gain Under Incomplete Information." *The Bell Journal of Economics* 9 (1):180–191. 16

Hyytiä, Esa, Aleksi Penttinen, and Reijo Sulonen. 2012. "Non-myopic vehicle and route selection in dynamic DARP with travel time and workload objectives." *Computers & Operations Research* 39 (12):3021–3030. 70

Itoh, Hideshi. 1991. "Incentives to Help in Multi-Agent Situations." *Econometrica* 59 (3):611–636. 47

Lazear, Edward. 1989. "Pay Equality and Industrial Politics." *Journal of Political Economy* 97 (3):561–580. 20

Levin, Jonathan and Steven Tadelis. 2005. "Profit Sharing and the Role of Professional Partnerships." *Quarterly Journal of Economics* 120 (1):131–171. 16

Ma, Hongyao, Fei Fang, and David C Parkes. 2018. "Spatio-Temporal Pricing for Ridesharing Platforms." *arXiv preprint arXiv:1801.04015* . 69

Madarász, Kristóf and Andrea Prat. 2017. "Sellers with Misspecified Models." *Review of Economic Studies* 84 (2):790–815. 16

Marku, Keler and Sergio Ocampo Diaz. 2017. "Robust Contracts in Common Agency." Tech. rep., University of Minnesota. 16

Monderer, D. and L. Shapley. 1996. "Potential Games." *Games and Economic Behavior* 14:124–143. 32

Raju, CVL, Y Narahari, and K Ravikumar. 2003. "Reinforcement learning applications in dynamic pricing of retail markets." In *E-Commerce, 2003. CEC 2003. IEEE International Conference on*. IEEE, 339–346. 70

Rana, Rupal and Fernando S Oliveira. 2014. "Real-time dynamic pricing in a non-stationary environment using model-free reinforcement learning." *Omega* 47:116–126. 70

Sannikov, Yuliy. 2008. "A continuous-time version of the principal-agent problem." *The Review of Economic Studies* 75 (3):957–984. 138

Yamashita, Takuro. 2015. "Implementation in Weakly Undominated Strategies: Optimality of Second-Price Auction and Posted-Price Mechanism." *Review of Economic Studies* 82 (3):1223–1246. 16

Yang, Jian, Patrick Jaillet, and Hani Mahmassani. 2004. "Real-time multivehicle truckload pickup and delivery problems." *Transportation Science* 38 (2):135–148. 70