Essays in Economic Theory

Inauguraldissertation

zur Erlangung des Grades eines Doktors
der Wirtschaftswissenschaften

durch

die Rechts- und Staatswissenschaftliche Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Sophie Alexandra Kreutzkamp

aus Bonn

2024

Dekan:	Prof. Dr. Jürgen von Hagen
Erstreferent:	Prof. Dr. Dezsö Szalay
Zweitreferent:	Prof. Dr. Benny Moldovanu
Tag der mündlichen Prüfung:	13. Juni 2023

Acknowledgements

First, I would like to thank my supervisors Dezsö Szalay and Benny Moldovanu for their invaluable support, continuous encouragement and guidance, and for many inspiring discussions throughout the PhD.

I also thank Philipp Strack for his advice and many valuable comments during my research visit at Yale and even after my return to Bonn. I am grateful to Dirk Bergemann, Francesc Dilmé, and Stephan Lauermann for many useful suggestions and feedback.

I was fortunate to have found two great coauthors, Axel Niemeyer and Finn Schmieter, for the second chapter of this dissertation. Working with them has been a very productive and joyful experience. Further, I would like to thank many more fellow students and colleagues, in particular Simon Block, Kailin Chen, Patrick Lahr, Justus Preusser, and Tobias Rachidi for making the last years a pleasant journey.

Finally, I am grateful to my family and friends, in particular to my parents, Sabine and Reinhold, for their constant support during the PhD.

Contents

Introduction

This dissertation consists of three self-contained essays in microeconomic theory and statistics. The first chapter contributes to the literature on information economics. It examines strategic information transmission in a sender-receiver game with endogenous learning. The second chapter on search and matching theory analyses a dynamic matching market in which match values increase over time, and agents have the option to rematch. The third chapter on cluster analysis deals with a constrained clustering model.

Chapter 1 contains my job market paper "Endogenous Information Acquisition in Cheap-Talk Games": It studies a communication game with endogenous learning: An expert publicly acquires costly private information about a state of nature and then communicates with a decision-maker by sending a cheap-talk message. Public learning (i) improves the communication outcome, and (ii) fundamentally changes the structure of communication: In Pareto efficient equilibria, the expert reveals all acquired information to the decision-maker. In general, the expert does not acquire full information—even if learning is costless. I provide a geometric characterization of the Pareto efficient equilibria for a general setting by identifying the extreme points for generic spaces of Blackwell experiments. These tools can also be applied to standard cheap-talk settings with a perfectly informed expert. Under posterior-mean preferences, any cheap-talk problem is solved by a convex combination of two *bi-pooling policies*. An application is the uniform-quadratic model, for which the best bi-pooling policies are characterized. Contrary to existing cheap-talk models, monotone partitions are not always optimal.

Chapter 2 builds on joint work with Axel Niemeyer and Finn Schmieter, "On-The-Match Search and Match-Specific Productivity Growth". We study a search-and-matching model with heterogeneous agents that continue to search on-the-match. In deciding with whom to match, agents must trade-off the flow utility provided by a partner against the stability of a match, i.e., the rate at which the partner leaves for another agent. Thus, stability determines and is determined by the agents' behavior, and consequently, there are multiple steady state equilibria. In almost every equilibrium, agents coordinate on payoff-dominated behavior. However, if there is match-specific productivity growth, i.e., if flow utility increases in the duration of a match, agents no longer fail to coordinate. We characterize the set of steady state equilibria that survive a perturbation with match-specific productivity growth. In some equilibria, less productive agents prefer to match with other less productive agents, suggesting an alternative explanation for assortative matching. In general, productivity growth can significantly alter equilibrium outcomes and sorting patterns: any match now becomes stable in the long run, but there is an incentive to foster growth in matches that are stable to begin with; either effect can dominate the agents' trade-off.

Chapter 3 is my third paper "Clustering with a Minimum Distance Constraint". This project studies the k-means clustering problem with one additional constraint: The distance between any two cluster centroids is bounded below by some constant. Specifying the minimum distance between cluster means determines the optimal number of clusters. I characterize analytical properties of the solution to the constrained clustering problem. The classical k-means clustering algorithm extended by the minimum distance constraint typically converges to an outcome that is a local solution, but cannot be the global solution to the clustering problem. Moreover, I propose a hypothesis test on uniformity of the underlying distribution of the original data based on the clustered data.

Chapter 1

Endogenous Information Acquisition in Cheap-Talk Games*

1.1 Introduction

Information plays a central role in many decision processes. Nevertheless, decision-makers often face a lack of information, preventing them from making good choices. Hence, effective information transmission is crucial. A simple, but functional form of information sharing is costless, non-binding and unverifiable communication: cheap talk. Communication is key in our daily life, and cheap talk matters: The success of big organizations depends on the interaction among its divisions. Another example of economic relevance is the communication strategy of central banks such as the Fed or ECB. This is cheap talk (see Duffy and Heinemann (2021)), and people care for what central banks say. So, their announcements have an effect on people's behavior, markets, and the economy as a whole.

This paper investigates a communication game with public learning. The literature on cheap-talk goes back to the influential work by Crawford and Sobel (1982): A biased sender transmits a message about a state to a receiver who takes a decision that affects both agents' payoffs. The underlying assumption in their model is that the sender is perfectly informed about the state. However, in many situations, even experts are not omniscient, but need to acquire information before they can give a recommendation. For example, central banks estimate some macroeconomic models before making announcements.

* I am grateful to Nemanja Antić, Andreas Asseyer, Sarah Auster, Dirk Bergemann, Alexander Bloedel, Andreas Blume, Nina Bobkova, Florian Brandl, Francesc Dilmé, Daniel Krähmer, Patrick Lahr, Stephan Lauermann, Benny Moldovanu, Harry Pei, Francisco Poggi, Francesco Squintani, Philipp Strack, Roland Strausz, and Dezső Szalay as well as audiences at Bonn and Yale, and at various seminars and conferences for valuable comments and suggestions. Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1 – 390838866 and through CRC TR 224 (Project B04) is gratefully acknowledged.

I deviate from Crawford and Sobel (1982) by assuming that the sender is not perfectly informed about the state. Instead, I allow the sender to decide which and how much information to acquire by conducting an experiment whose outcome yields an informative signal about the state. Information acquisition is flexible in the sense that the sender can choose an arbitrary statistical experiment à la Blackwell (1953) from a large set of available experiments.[1] Moreover, I study public learning, meaning that the receiver observes the choice of experiment, but not the outcome. In the central banks application, public learning means that the central banks disclose which macroeconomic models they use.[2]

In equilibrium, the receiver might remain partially uninformed about the state (i) due to a conflict of interest between the two parties, preventing the sender from sharing all acquired information with the receiver, and (ii) because the sender might not acquire all information if this is too costly. To better understand the overall effect and interplay of these two forces, the paper examines optimal experiments, i.e., Pareto efficient implementable experiments. There typically exist multiple optimal experiments because the two agents benefit from information asymmetrically: While the receiver is best off from the most informative experiments, the sender also has to bear the cost of learning so that the sender may be better off from experiments with limited information content. I provide an equilibrium selection criterion[3]—off-path coordination, meaning that agents coordinate on payoff-dominant behavior off the equilibrium path—that justifies the focus on Pareto efficient experiments. Besides, it is worthwhile to study optimal experiments even if one is not merely interested in Pareto efficient outcomes, but in the set of achievable equilibrium outcomes per se: Any equilibrium payoffs can be derived from convex combinations of the agents' payoffs in Pareto efficient and babbling equilibria.

My model is a general version of Crawford and Sobel (1982) going beyond single-peaked and single-crossing preferences. Apart from the usual continuity assumptions on utility and cost functions, and compactness assumptions on state and action spaces, I impose a monotonicity condition on the cost function, ensuring that the costs of an experiment are proportional to its Blackwell informativeness[4]. Cost functions commonly used in the literature on rational inattention, such as the entropy-

1. Flexible information acquisition is a standard assumption in the literature on information design (cf. Kamenica and Gentzkow (2011)) and coordination games (cf. Yang (2015)).

2. Another application of the model is communication based on survey data: A firm who contemplates launching a new product might contact a market researcher, who then conducts a survey among potential customers in order to evaluate the product's expected profitability. In this scenario, public learning could mean that the firm views the survey's questionnaire in advance to monitor the learning process to ensure that its outcome provides valuable information to the firm. Disclosure of the survey's result might not be possible due to data protection reasons.

3. For further selection criteria in the cheap-talk literature, see Chen, Kartik, and Sobel (2008), and Antić and Persico (2023).

4. An experiment is more informative in the sense of Blackwell (1953) than another experiment if the latter is a Blackwell garbling of the former. An experiment is most informative if no other available experiment is Blackwell more informative.

based cost function, satisfy this requirement. Also, free learning, i.e., zero cost, is a special case of the model.

I begin with a result that simplifies the identification of the set of optimal experiments: It is without loss of generality to restrict attention to equilibria in which the sender fully reveals all acquired information to the receiver. The intuition for this recommendation principle is straightforward: Since the receiver has the authority over decision making, the sender does not directly benefit from any information which is not transmitted to the receiver. It is more efficient if the sender only gathers information that the sender actually wants to forward to the receiver due to the cost caused by information gathering. Additionally, the paper provides an existence theorem for Pareto efficient equilibria in the general framework.

Having too much information can be harmful. To see this most clearly, consider the zero-cost case. Even if the sender has the option to choose a most informative experiment, the sender typically does not do so in an optimal equilibrium.[5] Since the sender cannot commit to fully revealing the experiment's outcome, the sender can transmit information to the receiver more credibly by having less private information: With less information, the sender has less incentives to misreport. In other words: The sender optimally acquires only information for which both agents agree on what is the best decision given that information, thus endogenously reducing the conflict of interest between the two parties.

I estimate the efficiency of the communication game's decision-making process by comparing the optimal equilibrium outcomes to Pareto efficient feasible outcomes. What are feasible outcomes? Suppose a social planner chooses an experiment in lieu of the sender and takes an action instead of the receiver based on the acquired information. An outcome that can be implemented this way is feasible. A common concern about cheap talk is that the lack of verifiability and commitment implies efficiency losses. It seems natural that frictions arise in the communication game if the two agents have unequal interests. But where exactly do those inefficiencies come from? Under certain regularity conditions, I show that a Pareto efficient feasible outcome is implementable if and only if there is a unique Pareto efficient feasible outcome. So if there is no disagreement about what is the best possible outcome, the two agents manage to coordinate in equilibrium to achieve this outcome. On the other hand, if the agents prefer different feasible outcomes, cheap talk involves inefficiencies. Consequently, different preferences per se do not lead to inefficiencies. Only if there is disagreement about best possible outcomes, frictions are inevitable under cheap talk. This finding provides an explanation why cheap talk is prevalent in real life.

5. This is an immediate consequence of the recommendation priciple. In cheap-talk settings with a perfectly informed sender, communication is typically coarse (cf. Crawford and Sobel (1982) and Green and Stokey (2007)).

Next, I geometrically characterize the optimal experiments for both the costless case and for concave[6] cost structures. I identify their support set of posterior beliefs. The cheap-talk problem is solved by an experiment that is a convex combination of up to two experiments. Their posterior beliefs satisfy an analogous condition as the convex independence condition of Crémer and McLean (1988, p.1251)). In the costless case, it is the notion of probabilistic independence as formulated by Lopomo, Rigotti, and Shannon (2022). For the costly case, a stronger version of this is needed: I define the notion of strong probabilistic independence. I show that the set of experiments satisfying this condition is the set of extreme points in the space of Blackwell experiments for arbitrary underlying distributions of the state. Intuitively, such experiments are optimal because they produce a certain expedient amount of information using the least possible number of different outcomes. If the state space is finite, the number of outcomes of an optimal experiment is bounded by twice the number of states.

An application are settings with a real-valued state space and posterior-mean preferences[7]: Optimal experiments are convex combinations of up to two bi-pooling policies—a term introduced by Arieli et al. (2023). A bi-pooling policy is a generalization of a monotone partition[8]: It divides the state space into convex subsets and associates up to two outcomes of the experiment to each subset. Similar results have been derived in the literature on Bayesian persuasion. This is an interesting observation on the relation between cheap talk and persuasion: While the two problems do not necessarily admit the same solution[9], they are equivalent in the sense that they admit a solution within the same class of experiments.

I apply the bi-pooling result to the uniform-quadratic case à la Crawford and Sobel (1982) with zero cost. Notably, the bi-pooling result provides sufficient structure to solve for the optimal experiment: The characterization builds on a proof by induction on the number of bi-pooling elements. The main qualitative finding—monotone partitions are not always optimal—is illustrated in Example 1.1:

Example 1.1. Figure 1.1 (a) shows the best equilibrium if the sender decides to become perfectly informed—it corresponds to the Pareto efficient Crawford-Sobel equilibrium. Can we do better by allowing the sender to acquire less than perfect

6. Concave cost include posterior separable cost—the standard assumption on cost in the literature on rational inattention (cf. Maćkowiak, Matějka, and Wiederholt (2018) and Matějka and McKay (2015)).

7. Such preferences do not depend on the whole distribution of conditional distribution of the state, but only on the distribution of its conditional expected value at any stage of the game.

8. A monotone partition divides the state space into convex subsets, and associates exactly one outcome of the experiment to each subset.

9. In general, optimal cheap talk and optimal Bayesian persuasion are not outcome-equivalent due to the commitment constraint that is imposed under cheap talk, but absent in the persuasion problem. Lipnowski (2020) derives outcome-equivalence of the two problems under certain conditions (namely finiteness of the action space and continuity of the sender's value function), which render the commitment constraint non-binding.

information? The answer is yes. Figure 1.1 (b) shows an equilibrium in which the sender only learns whether the state is in the first or second interval. So incentive compatibility means that the sender prefers the prescribed equilibrium action of each interval in expectation.[10] This equilibrium dominates the one in Figure 1.1 (a) because two equally informative messages are more efficient than one rather informative and one rather uninformative one. Figure 1.1 (c) shows an equilibrium that is a uniform monotone partition[11] with three intervals, which dominates the one in Figure 1.1 (b) because more messages are more efficient. It turns out that monotone partitions with four or more intervals are not implementable. Intuitively, incentive compatibility requires that the equilibrium actions must be sufficiently distant from one another. The partition in Figure 1.1 (d) is not incentive compatible because a_1 and a_2 as well as a_3 and a_4 are too close, respectively. However, incentive compatibility can be restored by adding a perturbation: By shifting the second interval further to the center, a_2 increases, and a_3 decreases. The partition in Figure 1.1 (e) is no longer monotone, but it is incentive compatible and dominates the best monotone partition of Figure 1.1 (c).

In general, the optimal experiment is either a uniform monotone partition, a non-uniform monotone partition with alternatingly sized intervals, or a bi-pooling policy with exactly two bi-pooling elements. This is an interesting finding because optimality of monotone partitions is prevalent in the uniform-quadratic case of related cheap-talk games (see Crawford and Sobel (1982), Pei (2015) and Ivanov (2010)). Comparative statics show that the outcome of my model is closest to the Pareto frontier of the set of feasible outcomes as compared to the related cheap-talk models (i.e., perfect information à la Crawford and Sobel (1982), private learning, public learning restricted to experiments being monotone partitions, mediation à la Krishna and Morgan (2004), long cheap talk à la Aumann and Hart (2003), noisy talk à la Blume, Board, and Kawamura (2007), etc.). This result suggests that public and flexible information acquisition is valuable. So the main practical takeaway is that agents should use these factors if they are available: In terms of the central banks application, this means, for instance, that it is beneficial for both the central banks and the population if central banks disclose how they collect their data. More abstractly speaking, the economic insight is that if two individuals with conflicting interests communicate, it is beneficial for both if the party that acquires information discloses how the information is gathered.

The cheap-talk model with public learning can also be interpreted as a persuasion model with partial commitment: The sender can commit to the experiment, but not to the message. So this case is in between full commitment, i.e, Bayesian

10. If the sender is perfectly informed, incentive compatibility requires that the sender prefers the prescribed equilibrium action at *every* state.

11. A uniform monotone partition is a monotone partition into equally sized intervals.

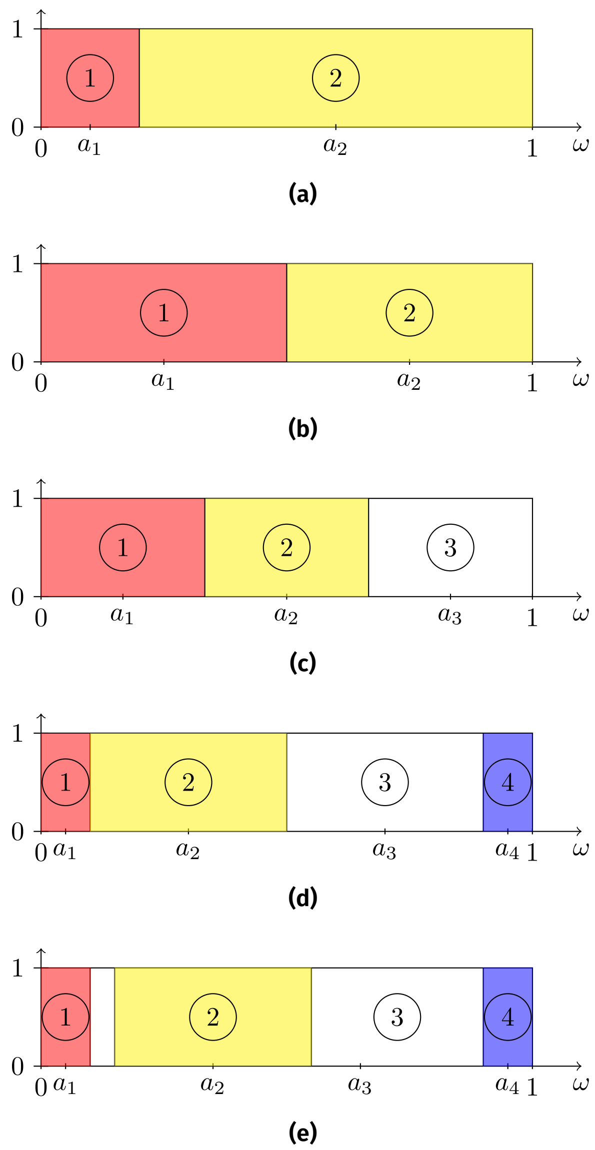


Figure 1.1. The graphs show (probabilistic) mappings from states ω to actions: Area i represents the mass of states associated with action a_i.

persuasion à la Kamenica and Gentzkow (2011), and no commitment (cheap talk with private learning). Is the outcome of partial commitment closer to no or full commitment? That is, is commitment to the information choice valuable on its own, or only in conjunction with commitment to the communication strategy? Comparative statics for the uniform quadratic model suggest that commitment with respect to the information choice has indeed an intrinsic value.

Notably, there is a structural equivalence between cheap talk with public learning and the standard cheap-talk setting with a perfectly informed sender. The tools provided in the present paper can also be applied to the standard setting. I find that the distribution over posterior beliefs in a Pareto efficient equilibrium satisfies the convex independence conditions, and I derive a version of the bi-pooling result for the setting with posterior-mean preferences. Besides, the number of messages in a Pareto efficient equilibrium can be bounded by twice the number of states.

Similarly, the results of this paper are also valid for standard persuasion settings as cheap talk with public learning is Bayesian persuasion with additional incentive constraints.

Related Literature. First, this paper contributes to the literature on Bayesian persuasion as it is a model with partial commitment. Following the pioneering work by Kamenica and Gentzkow (2011), the main tool of this literature is the concavification approach—the characterization of solutions via concave closures of the sender's valuation function.[12] This approach is also a useful technique to study cheap-talk games with state-independent sender preferences (see Lipnowski and Ravid (2020)) or settings with a binary state space (see Lyu and Suen (2022)). For general settings, the concavification approach is intractable due to a potentially high number of sender incentive compatibility constraints. Therefore, I use techniques from the literature on extreme points and majorization.

Since Kleiner, Moldovanu, and Strack's (2021) work on extreme points of monotonic functions under majorization constraints, the extreme-point approach—the characterization of solutions by extreme points—has become increasingly popular in the persuasion literature.[13] Kleiner, Moldovanu, and Strack (2021) and Arieli et al. (2023) show the optimality of bi-pooling policies in the standard persuasion problem with a real-valued state space and posterior-mean preferences.[14]

The present paper shows how the extreme-point technique can be adapted to cheap-talk settings with incentive compatibility constraints, and how they can be generalized beyond the case of real-valued state spaces and posterior-mean preferences.

Dworczak and Martini (2019) analyze optimal persuasion using a price-theoretic approach that yields sufficient conditions for a solution. My characterization via extreme points provides complementary necessary conditions for a solution.

Second, the present paper is related to the literature on strategic information transmission. Since the seminal work by Crawford and Sobel (1982), the embedment of endogenous learning into communication games has been an active subfield of the literature.

12. See Gentzkow and Kamenica (2014), Gentzkow and Kamenica (2016), and Ely (2017) for further applications of the concavification approach.

13. See Matysková and Montes (2021) and Candogan and Strack (2022) for further applications of the extreme-point approach in the persuasion literature.

14. In recent, independent work, Lou (2023) derives a version of the bi-pooling result in the cheap-talk game with public learning for a certain class of posterior-mean preferences using a convex programming approach, and also provides an analysis of the uniform-quadratic model.

Ivanov (2010) examines cheap talk under costless, public learning with one substantial difference: The receiver selects the experiment instead of the sender. He derives optimal monotone partitions for the uniform-quadratic case. I find that his model is equivalent to mine in terms of equilibrium outcomes. So as a byproduct, the tools in the present paper can also be applied to generalized versions of his model. In particular, they allow me to solve for the overall optimal partition in the uniform-quadratic case.

Pei (2015) discusses a cheap-talk game with costly, private learning. For the uniform-quadratic case, he shows that monotone partitions are chosen in equilibrium. The equilibrium outcomes in the limit case when costs converge to zero are equal to those in the model with a perfectly informed sender. This highlights the difference between public and private learning: The option for endogenous learning does not necessarily improve information transmission. It may be crucial that the sender can commit to the experiment.

Argenziano, Severinov, and Squintani (2016) investigate a communication game with both public and private costly information acquisition. The sender gathers information by repeatedly conducting a binary experiment. Its outcome is either a success or a failure. The probability for either outcome depends on the actual state realization. Each of these experiments can be converted into a Blackwell experiment, but the converse is not true. Therefore, the model predictions are different. For instance, the recommendation principle is not applicable.

Deimen and Szalay (2019) analyze public learning in a communication game with an endogenous bias: The conflict of interest between the two agents evolves through the learning process. In efficient equilibria, the sender acquires information that is equally beneficial for both agents. This reduces the endogenous bias, thus increasing cooperation between both parties. This relates to the intuition behind the implementability of Pareto efficient feasible outcomes in my model: Communication is best if the two agents are not biased towards different best feasible outcomes. This allows them to cooperate.

The present paper adds to this literature by investigating a general setting beyond single-peaked and single-crossing preferences or the uniform-quadratic case.

The paper is structured as follows: Section 1.2 introduces the model. Section 1.3 studies the implementable payoffs. Section 1.4 provides the geometric characterization of the optimal experiments. Section 1.5 deals with the application to bi-pooling policies. Section 1.6 discusses the relation to the standard cheap talk game. Section 1.7 contains the analysis of the uniform-quadratic case and comparative statics. Finally, Section 1.8 deals with the model variants, and Section 1.9 concludes.

1.2 Model

1.2.1 Setting

There is a sender S, a receiver R, and a state of the world ω following a distribution $\mu_0 \in \triangle\Omega$,[15], which is common knowledge. The state space Ω is convex.

The game proceeds as follows: First, the state realizes but remains unknown to both players. Second, the sender publicly chooses an experiment π, i.e., a distribution over posterior beliefs of the state $\mu \in \triangle\Omega$, at cost $c(\pi)$ and privately observes its outcome μ. The set of available experiments Π is a compact subset of $\triangle\left(\triangle\Omega\right)$. Third, the sender transmits a message $m \in \triangle\Omega$ to the receiver. Fourth, the receiver takes an action a from a compact, convex action space A. Payoffs are determined by the agents' von Neumann-Morgenstern utility functions $u_S : A \times \Omega \to \mathbb{R}$ and $u_R : A \times \Omega \to \mathbb{R}$, which are continuous on $A \times \Omega$.

Two assumptions on the learning cost and the set of experiments are imposed:[16]

Assumption 1.2.

- (*Monotonicity*): The cost function $c : \triangle\left(\triangle\Omega\right) \to \mathbb{R}$ is continuous and monotone, i.e., if π' is a Blackwell garbling of π, then $c\left(\pi\right) \geq c\left(\pi'\right)$.
- (*Richness*): If $\pi \in \Pi$ and π' is a Blackwell garbling of π, then $\pi' \in \Pi$.

Monotonicity captures the idea that Blackwell more precise information is costlier because information provided by a Blackwell garbling π' can also be generated by the original experiment π through appropriate mixing over its outcomes. Richness means that for any available experiment, the sender could also choose an arbitrary Blackwell garbling of it. Intuitively, the sender has enough flexibility in acquiring information about the state by having the option to choose any desired experiment—up to a certain level of precision.

1.2.2 Equilibrium Characterization

A sender strategy $\left(\sigma_{\mathscr{I}}, \sigma_{\mathscr{M}}\right)$ is an *information rule* $\sigma_{\mathscr{I}} \in \triangle\Pi$ and a *communication rule* $\sigma_{\mathscr{M}} : \Pi \times \triangle\Omega \to \triangle(\triangle\Omega)$. A receiver strategy is an *action rule* $\sigma_{\mathscr{A}} : \Pi \times \triangle\Omega \to \triangle A$. The sender's belief after choosing the experiment π and observing its outcome μ is μ. The receiver has a belief function $\mu_R : \Pi \times \triangle\Omega \to \triangle\Omega$ with $\mu_R\left(\cdot|\pi, m\right)$ indicating the receiver's belief after observing the sender's choice of the experiment π and receiving message m.

As is standard in the cheap-talk literature, I study perfect Bayesian equilibria.[17]

15. For a compact metrizable space X, let $\triangle X$ denote the set of distributions over X, endowed with the topology of weak convergence. For each $\chi \in \triangle X$, let $\operatorname{supp}(\chi)$ denote the support of χ.
16. These are generalizations of the assumptions by Pei (2015).
17. Restricting attention to PBE is for tractability: For any PBE, there exists an outcome-equivalent Bayesian equilibrium. Hence, all results in this paper also apply to the latter equilibrium concept.

Definition 1.3. A *perfect Bayesian equilibrium* (PBE) $E = \left\{\left(\left(\sigma_{\mathscr{I}}, \sigma_{\mathscr{M}}\right), \sigma_{\mathscr{A}}\right), \mu_R\right\}$ consists of a strategy profile $\left(\left(\sigma_{\mathscr{I}}, \sigma_{\mathscr{M}}\right), \sigma_{\mathscr{A}}\right)$ with beliefs μ_R such that

(i) the information rule $\sigma_{\mathscr{I}}$ is optimal given $\left(\sigma_{\mathscr{M}}, \sigma_{\mathscr{A}}\right)$, that is,

$$\operatorname{supp}(\sigma_{\mathscr{I}}) \subseteq \underset{\pi \in \Pi}{\arg\max} \iiiint u_S(a, \omega)\ d\sigma_{\mathscr{A}}(a|\pi, m)\ d\mu(\omega)\ d\sigma_{\mathscr{M}}\left(m|\pi, \mu\right) d\pi\left(\mu\right) - c(\pi)$$

(ii) the communication rule $\sigma_{\mathscr{M}}$ is optimal given $\sigma_{\mathscr{A}}$, i.e., for all $(\pi, \mu) \in \Pi \times \Delta\Omega$,

$$\operatorname{supp}\left(\sigma_{\mathscr{M}}\left(\pi, \mu\right)\right) \subseteq \underset{m \in \Delta\Omega}{\arg\max} \iint u_S(a, \omega)\ d\sigma_{\mathscr{A}}(a|\pi, m)\ d\mu(\omega)$$

(iii) the action rule $\sigma_{\mathscr{A}}$ is optimal given μ_R, that is, for all $(\pi, m) \in \Pi \times \Delta\Omega$,

$$\operatorname{supp}\left(\sigma_{\mathscr{A}}(\pi, m)\right) \subseteq \underset{a \in A}{\arg\max} \int u_R(a, \omega)\ d\mu_R(\omega|\pi, m)$$

(iv) the receiver's beliefs are derived from Bayes' rule, whenever possible.

The agents' ex-ante expected payoffs of an equilibrium E are $\mathscr{U}_S(E)$ and $\mathscr{U}_R(E)$.

1.3 Implementable Payoffs

Which payoffs can be attained in equilibrium? To evaluate this, let's first consider the best and worst possible outcomes:

Definition 1.4. A payoff profile $\left(\mathscr{U}_R, \mathscr{U}_S\right)$ is *implementable* if there exists a PBE E such that $\mathscr{U}_i = \mathscr{U}_i(E)$ for all $i \in \{S, R\}$. It is *best (worst) implementable* if there is no other implementable payoff profile $\left(\mathscr{U}_R', \mathscr{U}_S'\right)$ with $\mathscr{U}_i < \mathscr{U}_i'$ ($\mathscr{U}_i > \mathscr{U}_i'$) for all i.

The following lemma underpins the idea that for the sake of studying the set of implementable payoffs, it is sufficient to characterize the set of best and worst implementable payoffs: any implementable payoff profile can be represented as a convex combination of best and worst implementable payoff profiles.

Lemma 1.5. *Any implementable payoff profile lies in the convex hull of the set of the best and worst implementable payoff profiles.*

Figure 1.2 illustrates this argument: Any payoff profile outside area A is not implementable: For those in area B, at least one agent's payoff is smaller than their payoff in the worst implementable outcome X. If a payoff profile in area C, for instance Y, was implementable, it would be best implementable or there would exist another best implementable payoff profile Y' Pareto dominating Y. Payoffs in area D are not implementable as they Pareto dominate the best implementable payoffs.

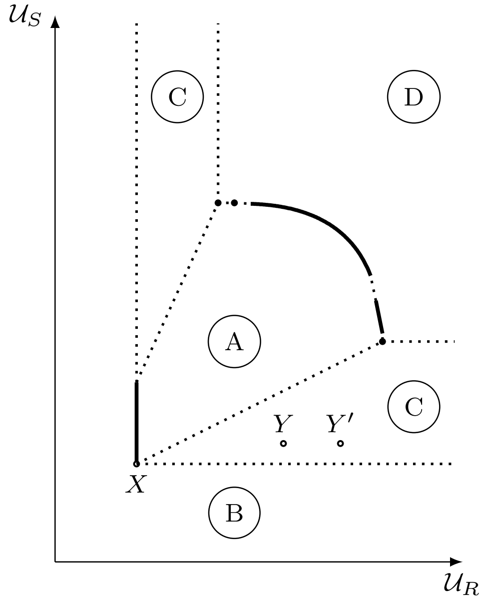


Figure 1.2. The solid line on the bottom left shows the worst implementable payoffs, and all other solid dots and lines show the best implementable payoffs. Area *A* is their convex hull.

1.3.1 Worst Implementable Payoffs

Worst implementable payoffs are attained in babbling equilibria, that is, PBE in which no information is transmitted to the receiver.[18]

Lemma 1.6. *Any worst implementable payoff profile can be generated in a babbling PBE. The set of worst implementable payoff profiles is non-empty, compact and convex.*

The existence result trivially follows from the existence of babbling equilibria. As illustrated in Figure 1.2, the receiver's payoff is constant across all babbling equilibria: The receiver achieves this payoff by choosing an optimal action given the prior distribution μ_0. Moreover, the receiver can secure this payoff by choosing that action in any PBE. Hence, it is the receiver's minimal payoff across all PBE. On the contrary, the sender's payoff is not necessarily the same across all babbling equilibria if the receiver's best response to the prior μ_0 is not unique. There is a sender-optimal and a sender-worst best response, and a range of payoffs for the sender can be generated through appropriate mixing over those two actions (see Figure 1.2). Given the receiver's strategy, the sender can secure any such payoff by choosing the uninformative experiment whose sole outcome is μ_0.

18. Formally, I define a babbling equilibrium as a PBE with $\mu_R(\cdot|\pi, m) = \mu_0$ for all $\pi, m \in \triangle\Omega$.

1.3.2 Pareto Efficient Implementable Payoffs

Pareto efficient implementable payoff profiles are a subset of the set best implementable payoff profiles. It is without loss of generality to focus on Pareto efficient implementable payoffs because the convex hull of the implementable payoffs in Lemma 1.5 can be derived from the set of Pareto efficient implementable payoffs. [19]

1.3.2.1 Recommendation Principle

This section provides a useful tool towards determining the set of Pareto efficient implementable payoff profiles: It is without loss of generality to focus on PBE where the sender (i) does not mix at the learning stage, and (ii) fully reveals the acquired information to the receiver.

Definition 1.7. An information rule $\sigma_{\mathscr{I}}$ is called *pure-strategy* if $|\operatorname{supp}(\sigma_{\mathscr{I}})| = 1$. A PBE is *fully revealing* if $\mu_R(\cdot|\pi, m) = \mu$ for every $m \in \operatorname{supp}(\sigma_{\mathscr{M}}(\pi, \mu))$, all $\mu \in \operatorname{supp}(\pi)$ and all $\pi \in \operatorname{supp}(\sigma_{\mathscr{I}})$. The communication rule in a fully revealing PBE is denoted by $\sigma_{\mathscr{M}}^{FR}$.

The following lemma states the recommendation principle:

Lemma 1.8. *Any Pareto efficient implementable payoff profile can be generated by a fully revealing PBE with a pure-strategy information rule. If c is strictly monotone[20], any PBE generating a Pareto efficient implementable payoff profile is fully revealing.*

The argument involves two steps. First, for any PBE with a mixed-strategy information rule, there exists a weakly Pareto dominant PBE with a pure-strategy information rule: The sender is indifferent between all experiments in the mixture, so one can choose the one that is best for the receiver to construct a PBE with a pure-strategy information rule. Second, for any PBE with a pure-strategy information rule, there exists a weakly Pareto dominant fully revealing PBE with a pure-strategy information rule: Consider a non-fully revealing PBE in which the sender chooses experiment π. Construct a new experiment π^* replicating the information that is transmitted in that PBE. Note that π^* is a Blackwell garbling of π, hence less costly. One can design a fully revealing PBE in which the sender chooses π^* all else equal. The receiver's payoff is the same and the sender's payoff is weakly higher.

The recommendation principle is interesting because (i) it is a preliminary result facilitating the characterization of Pareto efficient experiments, but more importantly (ii) it provides an insightful interpretation: Since information is costly, it is

19. To be precise, the convex hull of the implementable payoffs is equal to the convex hull of the worst implementable payoffs, the Pareto efficient implementable payoffs, and all implementable payoff profiles for which the sender's (receiver's) payoffs is the same as in the sender-optimal (receiver-optimal) equilibrium.

20. A cost function c is strictly monotone if $c(\pi) > c(\pi')$ whenever π' is a strict Blackwell garbling of π.

inefficient if the sender acquires information that is not transmitted to the receiver. Any additional private information is of no use to the sender. The receiver's decision can only be influenced via the revealed amount of information. Therefore, it is optimal if the sender only acquires information which both agents agree on what is the best decision, and which the sender can therefore truthfully share with the receiver.

Next, I prove the existence of Pareto efficient implementable payoff profiles.

Theorem 1.9. *The set of Pareto efficient implementable payoffs is non-empty and compact.*

Existence follows from an application of Berge's maximum theorem.[21] Assumption 1.2 and compactness of Π ensure that the corresponding optimization problem has a continuous objective function over a compact set. Let

$$\tilde{\mathscr{U}}_S\left(\pi,\tilde{\sigma}_{\mathscr{A}}\right) \equiv \int_{\triangle\Omega}\int_{\Omega}\int_{A} u_S(a,\omega)\ d\tilde{\sigma}_{\mathscr{A}}\left(a|\mu\right)\ d\mu(\omega)\ d\pi\left(\mu\right) - c(\pi)$$

$$\tilde{\mathscr{U}}_R\left(\pi,\tilde{\sigma}_{\mathscr{A}}\right) \equiv \int_{\triangle\Omega}\int_{\Omega}\int_{A} u_R(a,\omega)\ d\tilde{\sigma}_{\mathscr{A}}\left(a|\mu\right)\ d\mu(\omega)\ d\pi\left(\mu\right)$$

be the agents' ex-ante expected payoffs if the sender chooses experiment π, communicates truthfully, and the receiver chooses an action according to the strategy $\tilde{\sigma}_{\mathscr{A}} : \triangle\Omega \to \triangle A$. A payoff profile $\left(\mathscr{U}_S^*, \mathscr{U}_R^*\right)$ is a Pareto efficient implementable outcome if and only if there exists some π^* and some $\tilde{\sigma}^*_{\mathscr{A}}$ with $\mathscr{U}_i^* = \tilde{\mathscr{U}}_i\left(\pi^*, \tilde{\sigma}^*_{\mathscr{A}}\right)$ for each $i \in \{S, R\}$ solving

$$\max_{(\pi,\tilde{\sigma}_{\mathscr{A}})} \quad \tilde{\mathscr{U}}_S\left(\pi,\tilde{\sigma}_{\mathscr{A}}\right)$$

s.t.

$$\operatorname{supp}\left(\tilde{\sigma}_{\mathscr{A}}(\mu)\right) \subseteq \arg\max_{a\in A} \int u_R(a,\omega)\ d\mu(\omega) \quad \text{for all } \mu \in \triangle\Omega \tag{1.1}$$

$$\begin{aligned}&\iint u_S(a,\omega)\ d\tilde{\sigma}_{\mathscr{A}}\left(a|\mu\right)\ d\mu(\omega)\\ &\quad\geq \iint u_S(a,\omega)\ d\tilde{\sigma}_{\mathscr{A}}\left(a|\mu'\right)\ d\mu(\omega)\end{aligned} \quad \text{for all } \mu,\mu' \in \operatorname{supp}(\pi) \tag{1.2}$$

$$\tilde{\mathscr{U}}_S\left(\pi,\tilde{\sigma}_{\mathscr{A}}\right) \geq \mathscr{U}_S^0 \tag{1.3}$$

$$\tilde{\mathscr{U}}_R\left(\pi,\tilde{\sigma}_{\mathscr{A}}\right) \geq \mathscr{U}_R \tag{1.4}$$

for some $\mathscr{U}_R \geq \mathscr{U}_R^0$, where $\mathscr{U}_i^0$ denotes agent i's minimal ex-ante expected payoff in a PBE. This maximization problem determines the highest possible ex-ante expected payoff the sender can obtain in a fully revealing PBE with a pure-strategy information rule provided that the receiver's ex-ante expected payoff does not fall below a threshold value $\mathscr{U}_R$: The constraints (1.1) guarantee the optimality of the receiver's

21. See Aliprantis and Border (2006, p.570).

action rule. The inequality constraints (1.2) ensure that the sender has an incentive to report the experiment's observed outcome always truthfully. The participation constraint (1.3) makes sure that the sender is willing to acquire costly information, that is, the ex-ante expected payoff must not be lower than the minimal ex-ante expected payoff obtainable by acquiring no information. Condition (1.4) makes sure that the receiver's ex-ante expected payoff is at least $\mathcal{U}_R$.

The equilibrium implementing the best implementable payoffs $\left(\mathcal{U}_S^*, \mathcal{U}_R^*\right)$ is as follows: It is a fully revealing PBE with a pure-strategy information rule in which the sender chooses the experiment π^*. For any other experiment off the equilibrium path, the agents agree on a babbling outcome which yields the minimal ex-ante expected payoff $\mathcal{U}_S^0$ to the sender.

1.3.2.2 Off-Path Coordination

The literature on communication games focuses on sender- or receiver-optimal equilibria (cf. Kamenica and Gentzkow (2011), Gentzkow and Kamenica (2014), and Ambrus, Azevedo, and Kamada (2013)). So why study all Pareto efficient implementable payoff profiles beyond the reasons surrounding Lemma 1.5? A natural selection criterion is off-path coordination: The two players coordinate on Pareto efficient outcomes off the equilibrium path (when the sender chooses an experiment not prescribed by the information rule).

Definition 1.10. Let $\mathcal{P}$ ($\mathcal{P}^I$) be the set of Pareto efficient implementable (implementable) payoffs, and let $\mathcal{P}_i$ be the set of payoffs that agent i can attain in a Pareto efficient PBE. For $V \in \mathcal{P}_S$, let $\mathcal{U}_S^{\min}(V)$ be the minimal payoff the sender can attain in a fully revealing PBE by choosing the same experiment as in a fully revealing Pareto efficient PBE generating sender payoff V. A payoff profile $\left(\mathcal{U}_S, \mathcal{U}_R\right) \in \mathcal{P}^I$ satisfies *off-path coordination* if

$$\begin{aligned} \mathcal{U}_S &= \sup\left\{V \in \mathcal{P}_S : \mathcal{U}_S^{\min}(V) \geq V' \text{ for all } V' \in \mathcal{P}_S : V' \leq V\right\}, \\ \mathcal{U}_R &= \sup\left\{V : \left(\mathcal{U}_S, V\right) \in \mathcal{P}^I\right\}. \end{aligned} \tag{1.5}$$

Off-path coordination yields a unique outcome, which is Pareto efficient:

Lemma 1.11. *There exists a unique payoff profile $\left(\mathcal{U}_S^*, \mathcal{U}_R^*\right) \in \mathcal{P}^I$ satisfying off-path coordination. Moreover, $\left(\mathcal{U}_S^*, \mathcal{U}_R^*\right) \in \mathcal{P}$.*

If the receiver's best response in the action stage is always unique (which is the case under quadratic preferences, for instance), the model prediction under off-path coordination is the sender-optimal PBE. This is no longer true if the receiver is indifferent between several actions: The receiver could announce choosing an action from the indifference set being least favorable to the sender in case the sender deviates to an experiment off path. This is a credible threat. Therefore, off-path coordination typically yields an outcome in between the sender- and the receiver-optimum.

Analyzing all best implementable outcomes thus ensures to find the characteristics of the optimal experiment under this selection criterion.

Off-path coordination can be interpreted as a one-sided version of Antić and Persico's (2023) forward induction argument where the sender chooses an experiment at a cost instead of both agents choosing their preferences at a cost.

1.3.3 Feasible Payoffs

The set of Pareto efficient implementable payoffs depends on both the model's parameters (utility functions, set of available actions/experiments, etc.) and the game's structure (information acquisition and transmission by the sender, decision authority of the receiver): To better understand the interplay of these two forces, I compare the best possible outcomes that do not depend on the game's structure to the Pareto efficient implementable payoffs.

Suppose there is a social planner, i.e., a third neutral party, who chooses an experiment in lieu of the sender according to a mixed strategy $\bar{\sigma}_{\mathscr{I}} \in \triangle\Pi$ and, based on the experiment's outcome, an action in place of the receiver according to the strategy $\bar{\sigma}_{\mathscr{A}} : \Pi \times \triangle\Omega \to \triangle A$. Payoffs that can be generated via such an intervention by a social planner are of the form

$$\begin{aligned}
\bar{\mathscr{U}}_S\left(\bar{\sigma}_{\mathscr{I}}, \bar{\sigma}_{\mathscr{A}}\right) &\equiv \iiiint u_S(a,\omega)\ d\bar{\sigma}_{\mathscr{A}}(a|\pi,\mu)\ d\mu(\omega)\ d\pi\left(\mu\right) - c(\pi)\ d\bar{\sigma}_{\mathscr{I}}(\pi) \\
\bar{\mathscr{U}}_R\left(\bar{\sigma}_{\mathscr{I}}, \bar{\sigma}_{\mathscr{A}}\right) &\equiv \iiiint u_R(a,\omega)\ d\bar{\sigma}_{\mathscr{A}}(a|\pi,\mu)\ d\mu(\omega)\ d\pi\left(\mu\right)\ d\bar{\sigma}_{\mathscr{I}}(\pi).
\end{aligned}$$

Definition 1.12. A payoff profile $\left(\mathscr{U}_S, \mathscr{U}_R\right)$ is *feasible* if there exist some $\bar{\sigma}_{\mathscr{I}}$ and some $\bar{\sigma}_{\mathscr{A}}$ such that $\mathscr{U}_i = \bar{\mathscr{U}}_i\left(\bar{\sigma}_{\mathscr{I}}, \bar{\sigma}_{\mathscr{A}}\right)$ for all i.

The existence of Pareto efficient feasible payoff profiles is guaranteed:

Theorem 1.13. *The set of Pareto efficient feasible payoffs is non-empty and compact.*

The proof is analogous to Theorem 1.9. Payoffs $\left(\mathscr{U}_S^*, \mathscr{U}_R^*\right)$ are a Pareto efficient feasible outcome if and only if there is some $\left(\bar{\sigma}_{\mathscr{I}}^*, \bar{\sigma}_{\mathscr{A}}^*\right)$ with $\mathscr{U}_i^* = \bar{\mathscr{U}}_i\left(\bar{\sigma}_{\mathscr{I}}^*, \bar{\sigma}_{\mathscr{A}}^*\right)$ for all i solving

$$\begin{aligned}
&\max_{\left(\bar{\sigma}_{\mathscr{I}}, \bar{\sigma}_{\mathscr{A}}\right)} \quad \bar{\mathscr{U}}_S\left(\bar{\sigma}_{\mathscr{I}}, \bar{\sigma}_{\mathscr{A}}\right) \\
\text{s.t.} \quad &\bar{\mathscr{U}}_R\left(\bar{\sigma}_{\mathscr{I}}, \bar{\sigma}_{\mathscr{A}}\right) \geq \mathscr{U}_R. \qquad (1.6)
\end{aligned}$$

for some $\mathscr{U}_R \in \mathbb{R}$. While every implementable payoff profile is feasible, the converse is not true. So Pareto efficient implementable payoffs can, but need not be Pareto efficient feasible outcomes. The next theorem states a sufficient condition for the equivalence of Pareto efficient implementable and Pareto efficient feasible outcomes:

Theorem 1.14. *Suppose there exists a unique Pareto efficient feasible payoff profile. Then, it is the unique Pareto efficient implementable one.*

Uniqueness of the Pareto efficient feasible payoff profile implies that the sender- and receiver-optimal feasible payoff profile coincide. So both players agree on what is the best attainable outcome. This closes the gap between Pareto efficient feasible and Pareto efficient implementable payoffs because the two players have an incentive to coordinate in equilibrium to achieve the unique Pareto efficient feasible outcome.

Allowing for multiplicity annihilates the equivalence result:

Theorem 1.15. *Suppose there exist at least two Pareto efficient feasible outcomes. Then, the Pareto efficient feasible payoffs* $\left(\mathscr{U}_S, \mathscr{U}_R\right)$ *are not implementable if* $\operatorname{supp}(\sigma_{\mathscr{I}}) = \{\pi^{full}\}$ *for any* $\left(\bar{\sigma}_{\mathscr{I}}, \bar{\sigma}_{\mathscr{A}}\right)$ *with* $\mathscr{U}_i = \bar{\mathscr{U}}_i\left(\bar{\sigma}_{\mathscr{I}}, \bar{\sigma}_{\mathscr{A}}\right)$ *for each i.*

Consider first the receiver-optimal feasible payoff profile. Since it is feasible only with the fully informative experiment π^{full}, the sender would have an incentive to misreport in the communication stage to achieve the receiver-optimal feasible payoff profile, which is distinct by multiplicity. For all other Pareto efficient feasible payoffs, the receiver would have an incentive to choose another action rule after the sender fully revealed the outcome of π^{full} in order to generate the receiver-optimal feasible payoff profile instead.

To conclude, the frictions caused by information transmission from the acquirer to the decision-maker affect the Pareto efficient implementable outcomes beyond the model parameters whenever the sender and the receiver do not concur about the best achievable outcome, that is, whenever there are multiple Pareto efficient feasible payoffs.

The subsequent finding reinforces the idea that disagreement over optimal feasible outcomes leads to inefficiencies: It provides conditions under which uniqueness of Pareto efficient feasible payoffs is a necessary and sufficient condition for the equivalence of Pareto efficient implementability and feasibility. Let $A^*(\omega, u) = \arg\max_{a\in A} u\,(a, \omega)$ be the set of optimal actions if the state is ω given the utility function is u.

Corollary 1.16. *If* $A^*(\omega, \alpha u_S + (1-\alpha)u_R) \cap A^*(\omega', \alpha u_S + (1-\alpha)u_R) = \emptyset$ *for all* $\alpha \in [0,1]$ *and all* $\omega \neq \omega'$*, and* $c = 0$*, then a Pareto efficient feasible payoff profile is implementable if and only if it is unique.*

The set of Pareto efficient feasible and implementable payoffs either coincide or are disjoint—depending on whether the former one is a singleton or not. In that sense, the potential frictions introduced by the cheap-talk game influence all Pareto efficient implementable payoff profiles equally. The section concludes with an illustrating example:

Example 1.17. Consider the uniform-quadratic model à la Crawford and Sobel (1982) with an additive bias $b \in \left(0, \frac{1}{4}\right]$ and zero cost. Notice that $A^*(\omega, \alpha u_S + (1-\alpha)u_R) = \omega + (1-\alpha)b$ for all $\alpha, \omega \in [0,1]$. As illustrated in Figure 1.3, there is a continuum of Pareto efficient feasible payoff profiles none of which is implementable: The unique Pareto efficient implementable payoff profile $\left(\mathcal{U}_R^*, \mathcal{U}_S^*\right)$ is below the Pareto frontier of feasible payoffs.

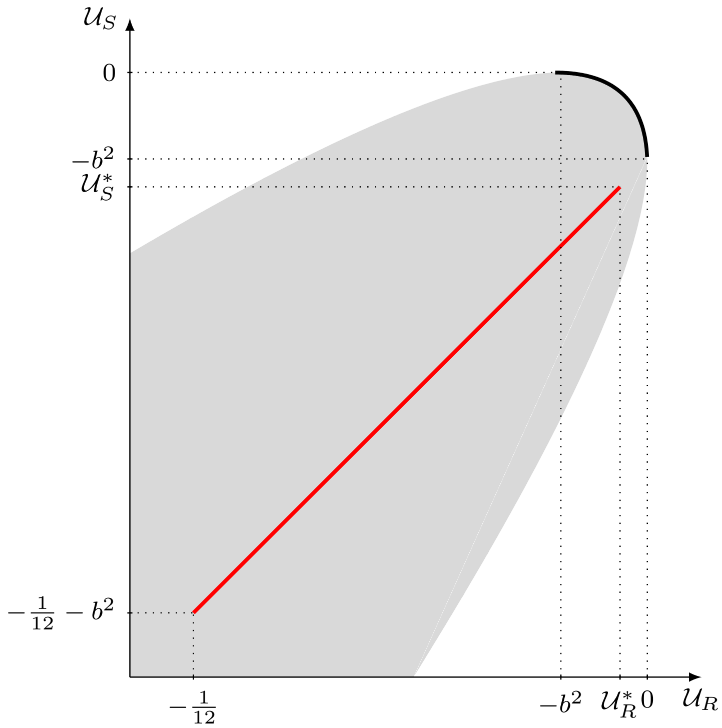


Figure 1.3. The shaded area represents the feasible payoffs, and the curve on the upper right is their Pareto frontier. The straight line shows the implementable payoffs.

1.4 Optimal Experiments

This section analyses optimal experiments, i.e., experiments that the sender chooses in a Pareto efficient PBE.

1.4.1 Costless Case

In the zero-cost case, optimal experiments have the following property:

Definition 1.18. An experiment π satisfies *probabilistic independence*[22] if

22. This follows the definition introduced by Lopomo, Rigotti, and Shannon (2022).

$$\text{for almost all } \mu \in \text{supp}(\pi) : \mu = \int_{\text{supp}(\pi)} \nu \, d\lambda(\nu) \quad \Rightarrow \quad \lambda = \delta_\mu \text{ a.e.,}$$

where δ_μ is the Dirac measure at μ.

For experiments with a finite support, probabilistic independence means that no posterior belief in the support can be represented as a convex combination of other posterior beliefs in the support. The following theorem formalizes the optimality result:

Theorem 1.19. *If $c = 0$, any Pareto efficient implementable payoff profile can be generated*[23] *by an experiment satisfying probabilistic independence.*

The proof is based on an extreme-point argument. Here is a sketch for the finite-support case: Consider an experiment π not satisfying probabilistic independence. There exists a posterior belief μ that can be represented as a convex combination of other posterior beliefs. Construct a new experiment π^* by replacing μ by the respective convex combination. Note that the π is a Blackwell garbling of π^*. Hence, the receiver's payoff is weakly higher in a fully revealing PBE in which the sender chooses π^* instead of π. Moreover, note that $\text{supp}(\pi^*) \subseteq \text{supp}(\pi)$. Therefore, if π is incentive compatible, so is π^*. Furthermore, the sender's payoff is weakly higher in a fully revealing PBE in which the sender chooses π^* instead of π by revealed preferences: If not, there must be a belief in the convex combination μ' so that the sender has an incentive to misreport μ when the true belief is μ'.

For binary state spaces, optimality of probabilistic independence implies that optimal experiments have at most two outcomes:

Corollary 1.20. *If $|\text{supp}(\mu_0)| = 2$, any Pareto efficient implementable payoff profile can be generated by an experiment π with $|\text{supp}(\pi)| \leq 2$.*

1.4.2 Costly Case

This section analyzes optimal experiments under the following additional assumption:

Assumption 1.21. The cost function c is concave.

Concave cost functions include affine functions, so in particular posterior separable cost functions[24]—the standard assumption on the cost structure in the literature on rational inattention (cf. Caplin and Dean (2013) or Caplin, Dean, and Leahy (2019)).

23. "Payoff profile X can be generated by experiment Y" is shortcut for "There exists a fully revealing PBE with a pure-strategy information rule in which the sender chooses experiment Y yielding ex-ante expected payoffs X."

24. A cost function c posterior separable if $c(\pi) = -k(\mu_0) + \int_{\Delta\Omega} k(\mu) \, d\pi(\mu)$ for all π for some convex function $k : \Delta\Omega \to \mathbb{R}$. Notice that any posterior separable cost function satisfies Assumption 1.2.

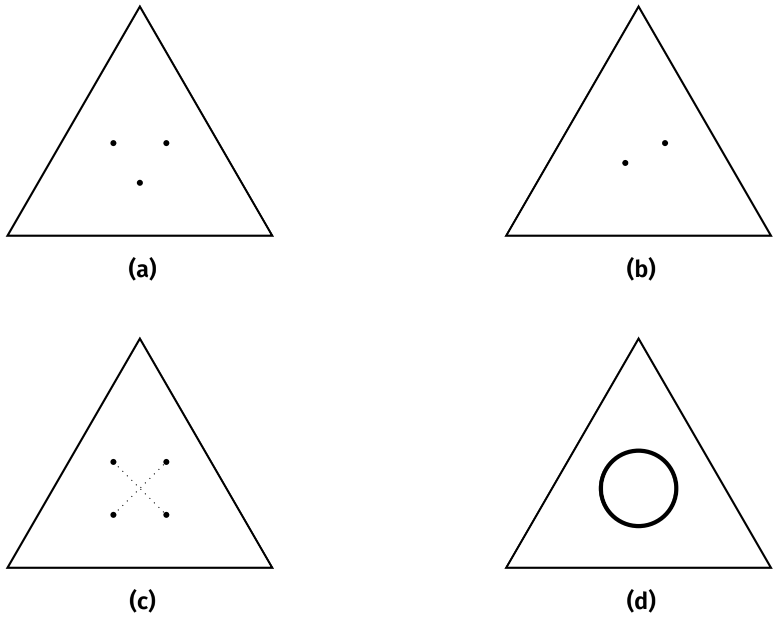


Figure 1.4. The 2-simplices show the convex hull of $\triangle\Omega$ for a ternary state space. The dots and lines in the simplices show the support of four different experiments.

To characterize optimal experiments in the costly case, I introduce a stronger notion than probabilistic independence:

Definition 1.22. An experiment π satisfies *strong probabilistic independence* if

$$\int_{\operatorname{supp}(\pi)} \nu \, d\lambda_1(\nu) = \int_{\operatorname{supp}(\pi)} \nu \, d\lambda_2(\nu) \quad \Rightarrow \quad \lambda_1 = \lambda_2 \text{ a.e..}$$

for all $\lambda_1, \lambda_2 \in \triangle\left(\triangle\Omega\right)$ with $p\lambda_i + (1-p)\lambda_i' = \pi$ for some $p \in (0,1)$ and $\lambda_i' \in \triangle\left(\triangle\Omega\right)$ for $i \in \{1,2\}$.

For experiments with a finite support, strong probabilistic independence means that no convex combination of posterior beliefs in the experiment's support can be represented by another convex combination of beliefs. Strong probabilistic independence implies probabilistic independence, but not vice versa. In fact, they are not equivalent. The following example illustrates the differences of the two concepts for a ternary state space:

Example 1.23. Figure 1.4 (a) shows an experiment with three different outcomes forming the edges of a triangle within the simplex—a subsimplex. Note that any belief in the subsimplex can be uniquely determined by a convex combination of the three edges. Hence, the experiment satisfies strong probabilistic independence. Figure 1.4 (b) shows an experiment with two outcomes forming the edges of a line—another subsimplex. Again, any belief on the line can be uniquely determined by a convex combination of the two edges. So the experiment satisfies strong probabilistic independence. Figure 1.4 (c) shows an experiment with four outcomes forming the edges of a rectangle. Note that no outcome can be represented as a convex combination of the other three outcomes, so the experiment satisfies probabilistic independence. However, it does not satisfy strong independence: The belief at the intersection of the two dotted lines can be represented by two different convex combinations of the experiment's outcomes. Indeed, this is a general pattern: No experiment with more than three outcomes satisfies strong probabilistic independence (cf. Corollary 1.26). On the contrary, this is not true for experiments satisfying probabilistic independence: They may have infinitely many different outcomes. Figure 1.4 (d) shows an experiment with a continuum of outcomes forming a circle. Obviously, this experiment does not satisfy strong probabilistic independence, but it does satisfy probabilistic independence: No outcome can be represented as a convex combination of other outcomes in the circle.

The geometric characterization of the example extends beyond the ternary-state case: Strong probabilistic independence means that the set of beliefs in the experiment's support represents the edges of a subsimplex within the simplex of all posterior beliefs $\triangle\Omega$.

Lemma 1.24. *An experiment is an extreme point of Π if and only if it satisfies strong probabilistic independence.*

Equipped with the notion of strong probabilistic independence, I can now state the optimality result for the costly case[25]: Optimal experiments are convex combinations of two experiments each of whom satisfies strong probabilistic independence.

Theorem 1.25. *Any Pareto efficient implementable payoff profile can be generated by an experiment of the form $\alpha\pi_1 + (1-\alpha)\pi_2$, where $\alpha \in [0,1]$, and π_1 and π_2 satisfy strong probabilistic independence. For sender- and receiver-optimal payoffs, it holds that $\alpha = 1$.*

The proof idea builds on another extreme-point argument as illustrated in Figure 1.5: Take an experiment π not satisfying the condition stated in Theorem 1.25. So in particular, π does not satisfy strong probabilistic independence. Hence, there exist two experiments π_1 and π_2 so that π corresponds to a convex combination of those two experiments, that is, π lies on the dotted line between π_1 and π_2. If either one, say π_2, does not satisfy strong probabilistic independence, it can be represented as a convex combination of two other experiments π_3 and π_4, so π_2 lies on the line between π_3 and π_4. But then, there is some convex combination of either

25. Concave cost include zero cost, so all results derived in this section also apply to the costless case, and indeed they are complementary: No result implies or is implied by another one.

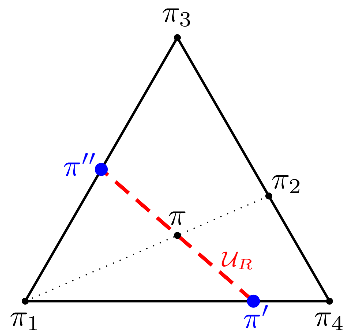


Figure 1.5. The 2-simplex shows the convex hull of the experiments π_1, π_3 and π_4. The dashed line between π' and π'' represents all experiments yielding payoff $\mathcal{U}_R$ to the receiver.

π_1 and π_3, π_1 and π_4, or π_3 and π_4 that is incentive compatible and Pareto dominates π: Incentive compatibility follows from the fact that the support of all three convex combinations is a subset of the support of π. Pareto dominance follows from the linearity of payoffs: By linearity of the receiver's payoff, the set of experiments that yield the same ex-ante expected payoff $\mathcal{U}_R$ to the receiver as the original experiment π form a straight line. By linearity of the sender's payoff, the sender's payoff is monotone along this line so that the optimum is attained at one of its intersections with the edges of the simplex, say π'. If π_1 and π_4 satisfy strong probabilistic independence, π' is a candidate solution. If not, one can repeat the process starting with π', forming a new simplex, etc. By compactness of Π, this algorithm either ends after finitely many steps, or it converges.

For finite state spaces, optimality of strong probabilistic independence implies that the size of an optimal experiment's support is bounded by twice the number of states.

Corollary 1.26. *If* $|\operatorname{supp}(\mu_0)| < \infty$*, any Pareto efficient implementable payoff profile can be generated by an experiment* π *with* $|\operatorname{supp}(\pi)| \leq 2|\operatorname{supp}(\mu_0)|$.

The proof idea is that any set of $|\Omega| + 1$ outcomes is linearly dependent, hence contradicts strong probabilistic independence. The number of outcomes is thus bounded by twice the number of states because the optimal experiment can be a convex combination of two experiments according to Theorem 1.25. Especially for environments with a small state space, Corollary 1.26 can be a helpful step towards finding an optimal experiment as it rules out a huge number of potential candidates for such an experiment. The section concludes with an application of the previous results.

Example 1.27. Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$, $A = \{a_1, a_2, a_3\}$, and $c = 0$. The agent's ordinal preferences are summarized in Table 1.1. If the sender is perfectly informed, it

Table 1.1. The entries in column 2 (3) show the receiver's (sender's) preferences given the state specified in column 1.

States	R	S
ω_1	$a_1 \succ a_2 \succ a_3$	$a_1 \succ a_2 \succ a_3$
ω_2	$a_2 \succ a_1 \succ a_3$	$a_3 \succ a_2 \succ a_1$
ω_3	$a_3 \succ a_1 \succ a_2$	$a_2 \succ a_3 \succ a_1$

is possible[26] that only babbling PBE exist: Too see this, note first that there is no PBE in which the receiver chooses all three actions with positive probability. If so, the sender would choose a message inducing action a_3 (a_2) if and only if the state is ω_2 (ω_3). Hence, both states would be fully revealed to the receiver who would then choose a_2 (a_3) if and only if the state is ω_2 (ω_3)—a contradiction. Second, there is no PBE in which the receiver chooses two actions with positive probability. Suppose there was one with a_1 and a_2.[27] The sender would induce action a_1 (a_2) if the state is ω_1 (ω_2 or ω_3). But if $u_R(a_2, \omega_3)$ is sufficiently small, the receiver prefers a_1 if the state is ω_2 or ω_3—a contradiction.

With public, flexible information acquisition, non-babbling PBE may exist as illustrated in Figure 1.6. First, there does not exist a PBE in which the receiver chooses all three actions: The sender would never truthfully report a belief for which the receiver chooses action a_3 as the sender prefers a_2 for any such belief (see Figure 1.6 (a) and (b)). But there exists an equilibrium in which the receiver chooses a_1 and a_2: There exist beliefs for which both agent's preferences over a_1 and a_2 coincide (the shaded areas in Figure 1.6 (d)) and whose convex hull includes μ_0. A candidate equilibrium is one generating beliefs μ_1 and μ_2.

The example provides an interesting insight on the value of endogenous learning: There are cases where no information can be transmitted if the sender is exogenously perfectly informed. Meaningful communication is only possible under endogenous learning.

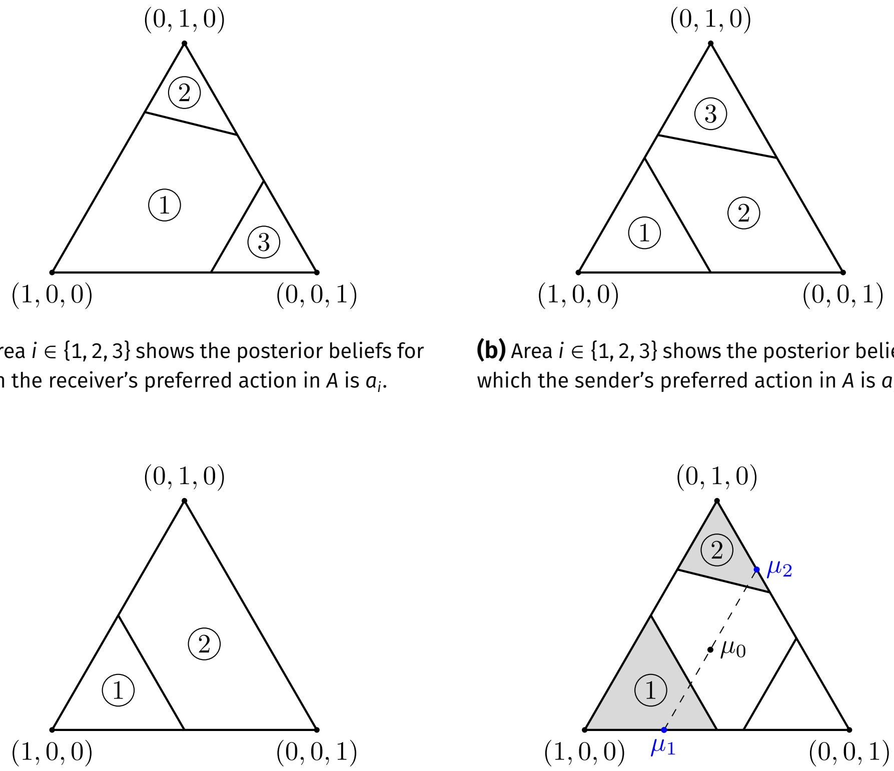
1.5 Optimality of Bi-Pooling Policies

This section studies optimal experiments for posterior-mean preferences.

Definition 1.28. A von Neumann-Morgenstern utility function u is *partially separable* if there are continuous functions $u_1 : A \to \mathbb{R}$, $u_2 : A \to \mathbb{R}$ and $u_3 : \Omega \to \mathbb{R}$ such that

$$u(a, \omega) = u_1(a) + u_2(a) \cdot \omega + u_3(\omega).$$

26. "Possible" means "There are cardinal preferences consistent with Table 1.1 for which only babbling PBE exist."
27. The argument for the other two action-combinations works analogously.



(a) Area $i \in \{1, 2, 3\}$ shows the posterior beliefs for which the receiver's preferred action in A is a_i.

(b) Area $i \in \{1, 2, 3\}$ shows the posterior beliefs for which the sender's preferred action in A is a_i.

(c) Area $i \in \{1, 2\}$ shows the posterior beliefs for which the sender's preferred action in $\{a_1, a_2\}$ is a_i.

(d) The shaded area $i \in \{1, 2\}$ shows the posterior beliefs for which the both agents' preferred action in $\{a_1, a_2\}$ is a_i. μ_1 and μ_2 represent the beliefs generated in an equilibrium.

Figure 1.6

Partially separable utility functions are additively separable in a and ω, except for one component that is linear in ω. This class of preferences contains commonly studied preferences such as quadratic utility[28] and quasi-linear utility[29].

The following results are derived under the assumption of a real-valued state and posterior-mean preferences and costs:

Assumption 1.29. Suppose that

- $\Omega \subset \mathbb{R}$,
- u_R and u_S are partially separable, and
- the cost function c satisfies

$$\rho_\pi = \rho_{\pi'} \quad \Rightarrow \quad c(\pi) = c(\pi')$$

28. See Crawford and Sobel (1982), Krishna and Morgan (2004), Ivanov (2010), Pei (2015), etc.
29. See Gentzkow and Kamenica (2016), Kolotilin et al. (2017), Candogan and Strack (2022), etc.

for all $\pi, \pi' \in \Pi$, where ρ_π denotes the distribution over posterior means of the state induced by the experiment π.

Posterior-mean costs include the class of Fréchet differentiable cost functions as introduced by Ravid, Roesler, and Szentes (2022), and in particular posterior separable cost functions that depend on the distribution of the posterior mean only:

Definition 1.30. A cost function c is *posterior-mean separable* if

$$c(\pi) = -k\left(\int_\Omega \omega \; d\mu_0(\omega)\right) + \int_{\Delta\Omega} k\left(\int_\Omega \omega \; d\mu(\omega)\right) d\pi(\mu).$$

for all π for some convex function $k : \mathbb{R} \to \mathbb{R}$.

Notice that any posterior-mean separable cost function satisfies Assumptions 1.2 and 1.21. Under these assumptions, both player's decisions at any stage of the game depend on the distribution of the posterior mean only. Moreover, both agents' payoffs are additively separable with respect to the prior μ_0 and the distribution over posteriors π.

Lemma 1.31. *Take any $\mu, \mu' \in \Delta\Omega$ with $\int_\Omega \omega \, d\mu(\omega) = \int_\Omega \omega \, d\mu'(\omega)$. Then, it holds that*

$$\arg\max_{a \in A} \int u_R(a, \omega) \; d\mu(\omega) = \arg\max_{a \in A} \int u_R(a, \omega) \; d\mu'(\omega),$$

and for any $a, a' \in A$,

$$\begin{aligned} &\int u_S(a, \omega) \; d\mu(\omega) \geq \int u_S(a, \omega) \; d\mu(\omega) \\ \Leftrightarrow &\int u_S(a, \omega) \; d\mu'(\omega) \geq \int u_S(a, \omega) \; d\mu'(\omega). \end{aligned}$$

For any $\sigma_{\mathscr{A}} : \Delta(\Delta\Omega) \times \Delta\Omega \to \Delta A$ and $\pi \in \Delta(\Delta\Omega)$, there exist functions $v_{i,1} : \Delta\Omega \to \mathbb{R}$ and $v_{i,2} : \Omega \to \mathbb{R}$ such that

$$\begin{aligned} &\iiint u_i(a, \omega) \; d\sigma_{\mathscr{A}}(a|\pi, \mu) \; d\mu(\omega) \; d\pi(\mu) \\ &\qquad = \int v_{i,1}(\mu) \; d\pi(\mu) + \int v_{i,2}(\omega) \; d\mu_0(\omega). \end{aligned} \tag{1.7}$$

Next, I introduce the notion of a bi-pooling policy[30], which is a generalization of a monotone partition: A monotone partition divides the state space into subintervals

30. The term "bi-pooling policy" is borrowed from Arieli et al. (2023). I extend their definition to distributions of the state with atoms.

assigning exactly one outcome of the experiment to each subinterval. A bi-pooling policy is a division of the state space into subintervals assigning up to two outcomes to each subinterval.

Definition 1.32. A *partitioning* of an experiment $\pi \in \Pi$ is a collection of closed intervals in Ω, denoted by $\{[\underline{\omega}_j, \bar{\omega}_j] | j \in J\}$, endowed with the corresponding collection of open intervals $\{(\underline{\omega}_j, \bar{\omega}_j) | j \in J\}$ such that

(1) $\bigcup_{j \in J} [\underline{\omega}_j, \bar{\omega}_j] = \Omega$,
(2) $(\underline{\omega}_j, \bar{\omega}_j) \cap (\underline{\omega}_{j'}, \bar{\omega}_{j'}) = \emptyset$ for all $j, j' \in J : j \neq j'$, and
(3) for all $\mu \in \text{supp}(\pi)$, there is exactly one $j \in J$ so that $\Pr\left(\tilde{\omega} \in [\underline{\omega}_j, \bar{\omega}_j] | \mu, \pi\right) = 1$.

A partitioning $\{[\underline{\omega}_j, \bar{\omega}_j] | j \in J\}$ of I is called *finest* if there is no $j^* \in J$ and $\omega^* \in [\underline{\omega}_{j^*}, \bar{\omega}_{j^*}]$ so that the collection $\{[\underline{\omega}_j, \bar{\omega}_j] | j \in J \backslash \{j^*\}\} \cup \{[\underline{\omega}_j, \omega^*], [\omega^*, \bar{\omega}_j]\}$ is a partitioning of π.

A partitioning is a collection of closed and open intervals so that the union of closed intervals covers the state space, the open intervals are pairwise disjoint, and each outcome μ is associated with one closed interval of the collection.[31] Endowed with the definition of a partitioning, bi-pooling policies can be specified:

Definition 1.33. Take an experiment $\pi \in \Pi$ with finest partitioning $\{[\underline{\omega}_j, \bar{\omega}_j] | j \in J\}$. Then, π is a *bi-pooling policy* if for any $j \in J$ with $(\underline{\omega}_j, \bar{\omega}_j) \cap \text{supp}(\mu_0) \neq \emptyset$, there are at most two outcomes $\mu, \mu' \in \text{supp}(\pi)$ that realize if $\omega \in \left(\underline{\omega}_j, \bar{\omega}_j\right)$. Two outcomes form a *2-partition* if they realize with positive probability on the same open interval $(\underline{\omega}_j, \bar{\omega}_j)$. All other outcomes form a *1-partition*.

The following corollary formalizes the optimality of bi-pooling policies:

Corollary 1.34. *Any Pareto efficient payoff profile can be generated by an experiment that is a convex combination of two bi-pooling policies. If $c = 0$, it can be generated by a bi-pooling policy.*

Bi-pooling policies satisfy strong probabilistic independence, but not every such experiment is a bi-pooling policy. However, the distributions over posterior means implied by bi-pooling policies form the extreme points of this space.[32] Intuitively, bi-pooling policies are optimal because they are the most informative experiments (they cannot be replicated by a mixture of other experiments) among all experiments for which both agents have a common interest in the revelation of their information content, that is, among all experiments the sender is willing to choose in a fully revealing PBE.

31. Each experiment has a partitioning: the singleton collection $\{\Omega\}$. Furthermore, the finest partitioning always exists and is unique.

32. More precisely, an experiment π is a bi-pooling policy if every experiment inducing the same distribution over posterior means as π satisfies probabilistic independence.

Remark 1. As shown by Kleiner, Moldovanu, and Strack (2021) and Arieli et al. (2023), bi-pooling policies are the optimal experiments in Bayesian-persuasion settings. Corollary 1.34 reveals a structural equivalence of cheap-talk and Bayesian persuasion: While the solution of the cheap-talk game is generally different from the solution under Bayesian persuasion, it belongs to the same class of experiments.

1.6 Cheap-Talk With a Perfectly Informed Sender

While the cheap-talk setting with public learning is conceptually different from the standard cheap-talk setting with a perfectly informed sender, there are two relations:

First, any equilibrium in the standard model is also an equilibrium under public learning (if the set of available experiments includes the fully informative one π^{full}): They can be constructed by letting the sender choose π^{full} on-path and imposing a babbling outcome everywhere off path. Consequently, Pareto efficient payoff profiles under public learning weakly dominate all implementable payoffs in the standard model.

Second and more importantly, the tools of the previous two sections can be applied to the standard model: Theorems 1.19 and 1.25, hence also Corollaries 1.20 and 1.26, continue to hold. To see this, note first that for any PBE E in the standard model, there exists an outcome-equivalent fully revealing PBE in the model with public learning and zero cost. In this fully revealing PBE E^*, the sender chooses an experiment π^* that replicates the outcome of the PBE E. Suppose π^* is not optimal in the model with public learning. Then, there exists an experiment π' with $\text{supp}(\pi') \subseteq \text{supp}(\pi^*)$ that the sender chooses in a Pareto dominant PBE in the model with public learning. This outcome is also implementable in the standard model as the incentive constraints are

$$\int u_S(a,\omega)d\sigma_{\mathscr{A}}(a|\mu) \geq \int u_S(a,\omega)d\sigma_{\mathscr{A}}(a|\mu') \quad \text{for all } \omega \in \text{supp}(\mu), \tag{1.8}$$

for all $\mu, \mu' \in \text{supp}(\pi')$, and thus a subset of those in the PBE E.

It is well known that in the standard model with a finite state space, Pareto efficient payoffs might not be implementable if the number of messages may not exceed the number of states. Corollary 1.26 sheds light on that: If the state space is finite, the number of messages in a Pareto efficient equilibrium can be bounded by twice the number of states.

A similar finding as the bi-pooling result can also be derived.

Definition 1.35. An experiment π is a *bi-partition* if for all $\Omega' \subseteq \text{supp}(\mu_0)$ with $\mu_0(\Omega') > 0$ and $|\Omega'| \geq 2$, it holds that

$$\bigcap_{\omega \in \Omega'} \{\bar{\omega}(\mu) \,:\, \omega \in \text{supp}(\mu) \text{ and } \text{Prob}(\mu|\omega) > 0\} \leq 2.$$

For any bi-pooling policy, there exists a bi-partition inducing the same distribution over posterior means of the state, but the converse does not hold true. For instance, all deterministic partitions, which map each state to exactly one posterior mean, are bi-partitions, but need not be bi-pooling policies.

Corollary 1.36. *Under Assumption 1.29, any Pareto efficient payoff profile can be generated by a convex combination of two bi-partitions. If $c = 0$, it can be generated by a bi-partition.*

The difference to Corollary 1.34 stems from the fact that the incentive constraints (1.8) depend on ω instead of μ.

1.7 The Uniform-Quadratic Case

This section characterizes the optimal experiments for the uniform-quadratic case à la Crawford and Sobel (1982) with zero cost: The state is uniformly distributed on the interval $[0,1]$. The agents' von Neumann-Morgenstern utility functions are $u_R(a,\omega) = -(a-\omega)^2$ and $u_S(a,\omega) = -(a-(\omega+b))^2$. Note that these utility functions are partially separable. By Corollary 1.34, optimal experiments are bi-pooling policies. They can be determined using the general maximization problem on page 15.

Lemma 1.37. *A tuple $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$ fulfills (1.1) and (1.2) if and only if $\operatorname{supp}\left(\sigma_{\mathscr{A}}(\mu)\right) = \{\int \omega \, d\mu(\omega)\}$ for all $\mu \in \operatorname{supp}(\pi)$ and $\left|\int \omega \, d\mu(\omega) - \int \omega \, d\mu'(\omega)\right| \geq 2b$ for all $\mu, \mu' \in \operatorname{supp}(\pi)$.*

Due to the quadratic preferences, the receiver's unique best response to outcome μ is the conditional mean $\int \omega \, d\mu(\omega)$ in a fully revealing PBE. The sender's incentive compatibility constraints reduce to the distance between any two induced posterior means of the state exceeding some constant, namely twice the bias.

Remark 2. In the standard model with a perfectly informed sender, incentive compatibility requires that the distance between any two induced adjacent actions is not constant, but increasing: Equilibria exhibit intervals of increasing length (cf. Crawford and Sobel (1982)). Where does the difference come from? In the model with public learning, the recommendation principle is responsible for the constant distance between induced action: This model can also be interpreted as one where the sender is perfectly informed, but the states of interest are the conditional means, i.e., the induced actions. These are in the center of the different intervals. In Crawford and Sobel (1982), intervals are of increasing length because the sender is inclined towards exaggerating on the right end of an interval, which are the states where the sender has the highest incentives to misreport. In my model, the sender's incentives towards exaggerating/undermining are equally distributed because the relevant state is not on the right end of an interval, but in the center.

The agents' ex-ante expected payoffs in a fully revealing PBE differ by a constant only: $\tilde{\mathscr{U}}_S\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) = \tilde{\mathscr{U}}_R\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) - b^2$ for all $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$ satisfying the conditions stated in Lemma 1.37. So there is a unique Pareto efficient implementable payoff profile, making (1.3) and (1.4) redundant.

The sender's incentive compatibility constraints imply that the optimal experiment has a finite support $\{\mu_1, \ldots, \mu_n\}$, where $n \in \mathbb{N}$. Define $\bar{\omega}_i \equiv \int \omega \, d\mu_i(\omega)$, and let p_i be the probability that the experiment's outcome is μ_i.

The optimal experiment depends on the bias: If $b > \frac{1}{4}$, no information is revealed.[33] Before deriving the exact structure of an optimal experiment for $b \leq \frac{1}{4}$, it is useful to think about its size n. The underlying trade-off is the following: Experiments with a larger n are more informative, thus payoff superior, but also less likely implementable in equilibrium. The best experiment is thus of a sufficiently large size n that is still consistent with incentive compatibility. The size of an optimal experiment can only take two different values:

Lemma 1.38. *If $b \in \left(\frac{1}{2n}, \frac{1}{2(n-1)}\right]$ for some $n \geq 3$, the optimal experiment is*

- *the uniform monotone partition of size $n-1$, or*
- *a bi-pooling policy of size n so that $\bar{\omega}_{i+1} - \bar{\omega}_i = 2b$ for all $i \in \{1, \ldots, n-1\}$, and both μ_1 and μ_n are 1-partitions.*

Neglecting all incentive compatibility constraints, the best experiment of a specific size is the uniform monotone partition of that size. Moreover, uniform monotone partitions of larger sizes dominate uniform monotone partitions of smaller sizes. Consequently, the uniform monotone partition of size $n-1$ dominates all experiments of size $n-1$ or smaller. Also, it satisfies incentive compatibility. Experiments of size $n+1$ or larger are not incentive compatible. The optimal experiment is thus either the uniform partition of size $n-1$, or a bi-pooling policy of size n.

1.7.1 Small Bias: $b \leq \frac{1}{12}$

For $b \in \left(\frac{1}{2n}, \frac{1}{2(n-1)}\right] \leq \frac{1}{12}$, the optimal bi-pooling policy of size n is either a non-uniform monotone partition whose 1-partitions are of alternating size, or a bi-pooling policy with exactly two 2-partitions—one between the second and third outcome, and another one between the second-last and third-last outcome.

Lemma 1.39. *Let $n \geq 7$ and $b \in \left(\frac{1}{2n}, \frac{1}{2(n-1)}\right]$. The optimal bi-pooling policy of size n satisfies $\bar{\omega}_i = \frac{1}{2} + b(2i - n - 1)$ for all i, and*

33. In this case, the distance between any two adjacent induced actions must be strictly larger than $\frac{1}{2}$, implying that at most two different actions on $[0, 1]$ can be implemented. However, the distance between these two actions may be no larger than $\frac{1}{2}$. Due to the uniform distribution of the state, the maximal distance $\frac{1}{2}$ can essentially be achieved if and only if the lower action is induced whenever $\omega \in (0, x)$ and the upper action is induced whenever $\omega \in (x, 1)$ for some $x \in (0, 1)$, that is, whenever the experiment is essentially a monotone partition.

- *if n is odd, there exists some $\hat{b}_n \in \left(\frac{1}{2n}, \frac{1}{2(n-1)}\right)$ so that*
  - *if $b \leq \hat{b}_n$, it is a monotone partition with*

$$p_i = \begin{cases} 1 - 2b(n-1) & \text{, if } i \text{ is odd} \\ 2b(n+1) - 1 & \text{, if } i \text{ is even} \end{cases} \tag{1.9}$$

  - *if $b \geq \hat{b}_n$, μ_2 and μ_3 as well as μ_{n-2} and μ_{n-1} form a 2-partition, respectively, and all other μ_i are 1-partitions such that*

$$p_i = \begin{cases} 1 - 2b(n-1) & \text{, if } i \in \{1, n\} \\ \frac{\left(2b(n^2-17)-(n-4)\right)\left(2b((n-2)n-7)-(n-4)\right)}{16b(n-4)^2} & \text{, if } i \in \{2, n-1\} \\ \frac{\left(2b(n^2-17)-(n-4)\right)\left((n-4)-2b(n-3)^2\right)}{16b(n-4)^2} & \text{, if } i \in \{3, n-2\} \\ \frac{2b(n-4+(-1)^i)}{n-4} & \text{, else} \end{cases} \tag{1.10}$$

- *if n is even, μ_2 and μ_3 as well as μ_{n-2} and μ_{n-1} form a 2-partition, respectively, and all other μ_i are 1-partitions such that*

$$p_i = \begin{cases} 1 - 2b(n-1) & \text{, if } i \in \{1, n\} \\ \frac{(2b(n+4)-1)(2b(n+2)-1)}{16b} & \text{, if } i \in \{2, n-1\} \\ \frac{(2b(n+4)-1)(1-2b(n-3))}{16b} & \text{, if } i \in \{3, n-2\} \\ 2b & \text{, else} \end{cases} \tag{1.11}$$

To give an intuition for this result, notice that uniform monotone partitions would be best, but are infeasible. Therefore, one has to give up on uniformity or monotonicity. If n is odd, one can keep monotonicity if the bias is small, but the partition has alternatingly sized 1-partitions. The discrepancy between even and odd 1-partitions increases as the bias increases. At some point, it is better to add 2-partitions to restore more similarly sized 1-partitions in the center of the state space. If n is even, the optimal experiment of size n cannot be a monotone partition.[34] The solution must contain 2-partitions. The pattern is similar to the odd case: To restore parity for most outcomes in the center, two 2-partitions close to the boundary of the state space are added.

Why is this optimal? First, note that 2-partitions involve a loss of information. Separating a 2-partition into two separate 1-partitions would be payoff dominant. However, adding a small number of 2-partitions (two due to symmetry) can increase

34. Any monotone partition for which all IC-constraints of adjacent induced actions are binding has alternatingly sized 1-partitions: $p_1 = p_3 = \ldots = p_{n-1}$ and $p_2 = p_4 = \ldots = p_n$. Moreover, the posterior mean $\bar{\omega}_i$ is always exactly at the center of the interval of states associated with μ_i. This implies $p_1 + p_2 = \ldots = p_{n-1} + p_n = 4b$ and thus $\sum_{i=1}^{n} p_i = \frac{n}{2} \cdot 4b = 2bn > 1$—a contradiction because $\sum_{i=1}^{n} p_i = 1$.

the efficiency of other outcomes. The first and last outcome form 1-partitions, so the 2-partitions cannot be directly at the margins of the state space. But why are they not further to the center? If so, there would be 1-partitions of alternating sizes at the margins, and less 1-partitions of more similar size in the center. This would be less efficient.

1.7.2 Large Bias: $\frac{1}{12} < b \leq \frac{1}{4}$

If the bias exceeds $\frac{1}{12}$, the optimal bi-pooling policy of maximal size n can be constructed straightforwardly. If $n = 3$, the optimal experiment is a monotone partition: Both the first and last outcome form a 1-partition by Lemma 1.38, so does the intermediate one. For $n = 4$, there are two options: The two intermediate outcomes either form a 2-partition or two separate 1-partitions. Recall that monotone partitions are not feasible because $n = 4$ is even. Consequently, the two intermediate outcomes must form a 2-partition. If $n = 5$, three different scenarios are possible: Either the second and third outcome form a 2-partition, or the third and fourth outcome form a 2-partition, or all outcomes form 1-partitions. Due to the symmetry of the uniform distribution, the optimal partition turns out to be symmetric. Hence, it is a monotone partition. For $n = 6$, one obtains by a similar symmetry-argument and by infeasibility of monotone partitions that both the second and the third as well as the fourth and the fifth outcome form a 2-partition, respectively.

Lemma 1.40. *Let $n \in \{3, 4, 5, 6\}$ and $b \in \left(\frac{1}{2n}, \frac{1}{2(n-1)}\right]$. The optimal bi-pooling policy of size n satisfies $\bar{\omega}_i = \frac{1}{2} + b(2i - n - 1)$ for all i, and*

- *if n is odd, it is a monotone partition with (1.9).*
- *if n is even, μ_2 and μ_3 as well as μ_{n-2} and μ_{n-1} form a 2-partition with (1.11).*

1.7.3 Globally Optimal Experiments

The underlying trade-off when comparing the optimal bi-pooling policy of size n with the optimal experiment of size $n - 1$, i.e., the uniform partition of that size, is the following: The optimal bi-pooling policy of size n has an additional outcome realization coming at the expense of non-uniform 1-partitions or 2-partitions to satisfy incentive compatibility. The latter effect dominates as the bias increases.

Theorem 1.41. *If $b \in \left(\frac{1}{2n}, \frac{1}{2(n-1)}\right]$ for some $n \geq 3$, there exists some $\bar{b}_n \in \left(\frac{1}{2n}, \frac{1}{2(n-1)}\right)$ such that the optimal experiment is the best bi-pooling policy of size n if $b \leq \bar{b}_n$ and it is the uniform partition of size $n - 1$ if $b \geq \bar{b}_n$. Furthermore, if $n \geq 7$ and n is odd, then $\bar{b}_n \geq \hat{b}_n$.*

1.7.4 Comparative Statics

This section compares the outcome of the model with public learning to the outcome of other communication games.

First, consider the cheap-talk game with a perfectly informed sender à la Crawford and Sobel (1982). For any $b \in (0, \frac{1}{4})$, the Pareto efficient implementable outcome in this model is strictly dominated by the best outcome under public learning (cf. Proposition 1.42).

Second, consider cheap talk with private learning à la Pei (2015). In fact, the best outcome with private learning is equivalent to the best Crawford-Sobel equilibrium, and thus strictly dominated by the best outcome under public learning. Why is this? Under private learning (and zero cost), the sender chooses the fully informative experiment in every equilibrium, thus becoming perfectly informed. So, private learning has no effect, while public learning has a significant effect.

Third, consider cheap talk with public learning à la Ivanov (2010): Information acquisition is not flexible, but the sender can only choose monotone partitions. By Theorem 1.41, Ivanov's (2010) solution is the globally optimal experiment if and only if the latter is a monotone partition. But if it is a bi-pooling policy with proper 2-partitions, the difference can be substantial, as illustrated in Figure 1.7: The solution under public, flexible learning lies mid-way between the solution under public, restricted learning and the Pareto frontier of feasible payoffs. So, flexible information acquisition—the option to choose non-monotone partitions—is valuable. All

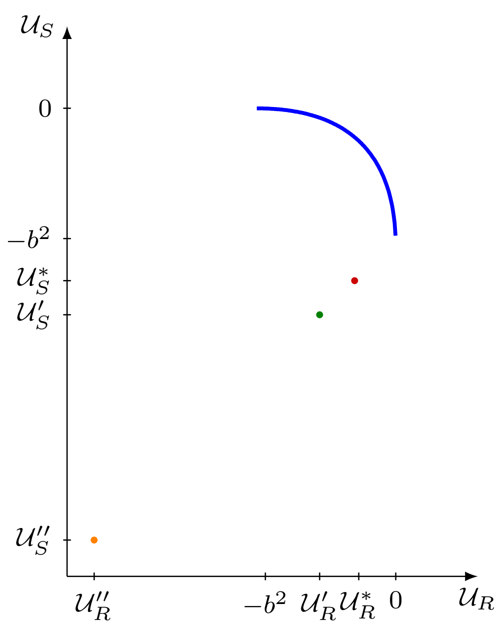


Figure 1.7. The curve represents the Pareto frontier of feasible payoffs. The dot at $(\mathcal{U}_S^*, \mathcal{U}_R^*)$ shows the optimal payoffs under public learning, $(\mathcal{U}_S', \mathcal{U}_R')$ those under public learning among the set of monotone partitions, and $(\mathcal{U}_S'', \mathcal{U}_R'')$ those under perfect information/private learning.

in all, the figure illustrates that the best outcome under public, flexible learning is closest to the Pareto frontier of feasible payoffs.

How does public learning compare to Bayesian persuasion à la Kamenica and Gentzkow (2011)? In the uniform-quadratic setting, the sender commits to the fully informative experiment, and truthfully reports it to the receiver under optimal persuasion. Payoffs are $(\mathcal{U}_S^{BP}, \mathcal{U}_R^{BP}) = (-b^2, 0)$—the right endpoint of the Pareto frontier. Cheap talk with public learning is a model of partial commitment (commitment to information, but not to communication), and Crawford and Sobel (1982) can be interpreted as an environment with no commitment. The payoffs under partial commitment are closer to those under full commitment than to those under no commitment:

Proposition 1.42. *For any $b \in \left(0, \frac{1}{4}\right)$, it holds that $\mathcal{U}_i^{BP} - \mathcal{U}_i^* \leq \mathcal{U}_i^* - \mathcal{U}_i''$ for all i.*

The value of commitment to information is higher than the value of commitment to communication: The efficiency loss of giving up on the latter commitment device while keeping the former one is not too large. This is good news to the information design literature: A common critique about Bayesian persuasion is that commitment to communication is hard to apply to real-world settings. On the other hand, there are applications for which commitment to information makes sense (see the example of survey data from the introduction). Since the latter commitment device dominates the former (in terms of value), the model of partial commitment could be an adequate alternative to Bayesian persuasion.

Finally, consider mediation à la Krishna and Morgan (2004): Suppose the sender is perfectly informed about the state, but does not directly communicate with the receiver. Instead, the sender transmits a message to a neutral mediator who then talks to the receiver. Can better equilibrium outcomes be attained than under public learning? In general, public learning and mediation cannot be ranked: The sender's incentives to report information are affected at different stages: Under public learning, the sender has less incentives to misreport by deciding to acquire less than perfect information. On the other hand, the mediator can enforce outcomes that are beneficial to both agents: If the sender communicates with the receiver directly, the sender would always send a message that implements the best possible action among all that the receiver chooses in equilibrium given the sender's information about the state. A mediator is not restricted to that (see the mediation solution by Krishna and Morgan (2004), for instance). In the uniform-quadratic case, public learning outperforms mediation:

Proposition 1.43. *For any $b \in \left(0, \frac{1}{2}\right)$, the unique Pareto efficient implementable payoff profile under public learning dominates any payoff profile under mediation.*

Public learning also outperforms long cheap talk à la Aumann and Hart (2003) as mediation dominates long cheap talk (see Krishna and Morgan (2004)), and it dominates noisy talk à la Blume, Board, and Kawamura (2007) as mediation and optimal noisy talk are payoff-equivalent.

Note that delegation à la Dessein (2002) is not Pareto comparable to public learning: The delegation solution corresponds to the sender-optimal feasible payoff profile: the left endpoint of the Pareto frontier in Figure 1.7. The receiver is strictly worse off under delegation for any $b \in \left(0, \frac{1}{4}\right)$ (cf. Ivanov (2010)) than under public learning, but the sender is not.

1.8 Model Variants

1.8.1 Experiment Choice by the Receiver

Consider the following model variant: First, the receiver publicly chooses an experiment π. Then, the sender decides whether to accept or reject π. If the sender accepts, the sender pays the cost $c(\pi)$, privately observes the experiment's outcome and sends a cheap-talk message to the receiver who then takes an action. If the sender rejects, communication breaks down and the receiver chooses an action based on the prior belief.[35] The recommendation principle remains valid: The Pareto efficient implementable payoff profiles of the original model can be generated by a fully revealing PBE in which the sender chooses a pure-strategy information rule, and a babbling outcome is implemented everywhere off-path. Since both agents' on-path payoffs exceed their payoffs under the babbling outcome, it does not matter whether the sender or the receiver chooses the experiment. Consequently, the model predictions do not change:

Proposition 1.44. *The set of Pareto efficient implementable payoff profiles in the model where the receiver chooses the experiment coincides with the one of the original model.*

1.8.2 Private Learning

Suppose now the sender chooses the experiment privately instead of publicly, that is, the sender cannot commit to the choice of the experiment. A version of recommendation principle can be established (cf. Pei (2015)). Any PBE of the model variant is a PBE of the original model, but the converse does not hold true: If the sender chooses an experiment in absence of commitment, the sender will also choose it with commitment power.

Proposition 1.45. *Any Pareto efficient implementable payoff profile under private learning is dominated by an implementable payoff profile under public learning.*

35. An alternative interpretation of rejection is that the sender chooses the uninformative experiment, and the communication stage yields a babbling outcome.

1.8.3 Bayesian Persuasion

Under Bayesian persuasion, the sender can commit to both the experiment and the full revelation of its outcome. Therefore, higher payoffs can be achieved in equilibrium:

Proposition 1.46. *Any Pareto efficient implementable payoff profile under public learning is dominated by an implementable payoff profile under Bayesian persuasion.*

1.9 Conclusion

This paper explores the effects of costly information procurement in a model of strategic information transmission from an expert to a decision-maker. Despite a conflict of interest between the two parties, full communication is attained in Pareto efficient equilibria. Less information acquisition leads to more credible communication, and thus more information transmission. The main contribution is a geometric characterization of the optimal experiments. I identify the set of extreme points in the space of Blackwell experiments. It is (i) a helpful tool for applications, and (ii) a robust pattern of information aggregation in different communication settings. The results carry over to the standard cheap-talk setting with a perfectly informed sender and Bayesian persuasion, for instance. It can therefore be considered as a promising approach to study more general heterogeneous-agents models with asymmetric information.

Appendix 1.A Proofs: Implementable Payoffs

Proof of Lemma 1.5. The proof directly follows from the description in the main text.

1.A.1 Section 1.3.1

Proof of Lemma 1.6. First, I show that the set of babbling equilibria $\mathcal{E}_0$ is non-empty by explicitly constructing one: Since Π is non-empty (by implicit assumption), and π^0 is a Blackwell garbling of any experiment, it follows from Assumption 1.2 that $\pi^0 \in \Pi$ and $c\left(\pi^0\right) = \min_{\pi \in \Pi} c(\pi)$. Fix some $m_0 \in \triangle\Omega$ and some $a_0 \in \arg\min_{a \in A_0} \int u_S(a, \omega)\, d\mu_0(\omega)$ where $A_0 \equiv \arg\max_{a \in A} \int u_R(a, \omega)\, d\mu_0(\omega)$. Notice that a_0 is well-defined: Since $A \times \Omega$ is compact and u_R is continuous on $A \times \Omega$, u_R is bounded on $A \times \Omega$. By the Dominated Convergence Theorem, continuity of $u_R(\cdot, \omega)$ on A for each $\omega \in \Omega$ implies continuity of $\int u_R(\cdot, \omega)\, d\mu_0(\omega)$ on A. Consequently, $\int u_R(\cdot, \omega)\, d\mu_0(\omega)$ attains a maximum on the compact set A, i.e., A_0 is non-empty, by the Weierstrass extreme value theorem. Consider the strategy profile $\left(\left(\sigma^0_{\mathscr{I}}, \sigma^0_{\mathscr{M}}\right), \sigma^0_{\mathscr{A}}\right)$ and beliefs μ^0_R with

$$
\begin{aligned}
\operatorname{supp}\left(\sigma^0_{\mathscr{I}}\right) &= \{\pi^0\}, \\
\operatorname{supp}\left(\sigma^0_{\mathscr{M}}(\pi,\mu)\right) &= \{m_0\} \quad \text{for all } \mu \in \Delta\Omega \text{ and all } \pi \in \Pi, \\
\operatorname{supp}\left(\sigma^0_{\mathscr{A}}(\pi,m)\right) &= \{a_0\} \quad \text{for all } m \in \Delta\Omega \text{ and all } \pi \in \Pi, \text{ and} \\
\mu^0_R(\cdot|\pi,m) &= \mu_0 \quad \text{for all } m \in \Delta\Omega \text{ and all } \pi \in \Pi.
\end{aligned}
$$

It can be easily verified that these profiles constitute a PBE E^0: First, the receiver's beliefs are always consistent with Bayes' rule on the equilibrium path, and it is μ_0 in any subgame in the action stage. Hence, it is—by definition of a_0—optimal for her to choose a_0 always. Likewise, it is a best response for the sender to transmit the uninformative message m_0 always because the receiver does not condition her action decision on the sender's message. Third, note that both agents' expected utilities (excluding the sender's cost) are constant across all subgames after the information acquisition stage. Therefore, the sender optimally chooses an experiment which induces the lowest possible cost in the information acquisition stage. Since $c\left(\pi^0\right) = \min_{\pi\in\Pi} c(\pi)$, it is optimal to take experiment π^0. By definition of μ^0_R, E^0 is a babbling equilibrium.

Second, the receiver's payoff is constant across all babbling equilibria by optimality of the receiver's action rule. Moreover, the receiver can secure this payoff $\mathscr{U}_R(E^0)$ in any PBE by choosing the action rule $\sigma^0_{\mathscr{A}}$. The sender can secure $\mathscr{U}_S(E^0)$ in any PBE by choosing the information rule $\sigma^0_{\mathscr{I}}$. However, the sender's payoff is not necessarily constant across all babbling equilibria. Depending on the receiver's action rule, the sender's payoff in a babbling PBE can take any value in $\left[\inf_{E\in\mathscr{E}_0} \mathscr{U}_S(E), \sup_{E\in\mathscr{E}_0} \mathscr{U}_S(E)\right]$. The set of worst implementable payoff profiles is $\left\{\left(\mathscr{U}_R(E^0), \mathscr{U}_S\right) | \mathscr{U}_S \in \left[\inf_{E\in\mathscr{E}_0} \mathscr{U}_S(E), \sup_{E\in\mathscr{E}_0} \mathscr{U}_S(E)\right]\right\}$, which completes the proof.

1.A.2 Section 1.3.2

Proof of Lemma 1.8. First, I show that any PBE with a mixed-strategy information rule is Pareto dominated by a PBE with a pure-strategy information rule: Fix a PBE $E = \left\{\left(\left(\sigma_{\mathscr{I}}, \sigma_{\mathscr{M}}\right), \sigma_{\mathscr{A}}\right), \mu_R\right\}$. For any $\pi \in \operatorname{supp}(\sigma_{\mathscr{I}})$, let $E^\pi = \left\{\left(\left(\sigma^\pi_{\mathscr{I}}, \sigma_{\mathscr{M}}\right), \sigma_{\mathscr{A}}\right), \mu_R\right\}$ where $\operatorname{supp}(\sigma^\pi_{\mathscr{I}}) = \{\pi\}$, and note that E^π is a PBE, too. By optimality of $\sigma_{\mathscr{I}}$ for the sender in the PBE E, it follows for all $\pi, \pi' \in \operatorname{supp}(\sigma_{\mathscr{I}})$ that $\mathscr{U}_S(E^\pi) = \mathscr{U}_S\left(E^{\pi'}\right)$ because

$$
\begin{aligned}
&\iiiint u_S(a,\omega)\; d\sigma_{\mathscr{A}}(a|\pi,m)\; d\mu(\omega)\; d\sigma_{\mathscr{M}}\left(m|\pi,\mu\right)\; d\pi\left(\mu\right) - c(\pi) \\
&\quad = \iiiint u_S(a,\omega)\; d\sigma_{\mathscr{A}}\left(a|\pi',m\right)\; d\mu(\omega)\; d\sigma_{\mathscr{M}}\left(m|\pi',\mu\right)\; d\pi'\left(\mu\right) - c\left(\pi'\right).
\end{aligned}
$$

Take $\pi \in \operatorname{supp}(\sigma_{\mathscr{I}})$ with $\mathscr{U}_R(E^\pi) \geq \mathscr{U}_R(E)$, which is possible because $\mathscr{U}_R(E) = \int_\Pi \mathscr{U}_R(E^\pi)\; d\sigma_{\mathscr{I}}(\pi)$. The PBE E^π generates payoffs $\left(\mathscr{U}_S(E), \mathscr{U}_R(E^\pi)\right)$, so E^π weakly dominates E.

Second, any PBE with a pure-strategy information rule is dominated by a fully revealing PBE with a pure-strategy information rule: Fix a PBE E with $\text{supp}(\sigma_{\mathscr{I}}) = \{\pi^*\}$ for some $\pi^* \in \Pi$. Consider the experiment π^{**} with outcomes $\{\mu_R(\cdot|\pi^*, m)\}_{m \in \bigcup_{\mu \in \text{supp}(\pi^*)} \text{supp}(\sigma_{\mathscr{M}}(\pi^*, \mu))}$ distributed according to $\int_{\mu \in \text{supp}(\pi^*)} \sigma_{\mathscr{M}}(\cdot|\pi^*, \mu)\, d\pi^*(\mu)$. In words, the distribution over outcomes under the experiment π^{**} corresponds to the distribution over the receiver's on-path beliefs in the PBE E. By Assumption 1.2, $\pi^{**} \in \Pi$ as π^{**} is a Blackwell garbling of π^*, and $c(\pi^{**}) \leq c(\pi^*)$. One can find a fully revealing PBE E' in which the sender chooses π^{**}. The agents' expected utilities are the same as in the PBE E, but the sender's cost are smaller in E'. Hence, E' dominates E. In order to construct the fully revealing PBE, consider the strategy profile $\left(\left(\sigma'_{\mathscr{I}}, \sigma'_{\mathscr{M}}\right), \sigma'_{\mathscr{A}}\right)$ and beliefs μ'_R with $\text{supp}(\sigma'_{\mathscr{I}}) = \{\pi^{**}\}$,

$$\sigma'_{\mathscr{M}}(\cdot|\pi, \mu) = \begin{cases} \sigma^{\text{FR}}_{\mathscr{M}}(\cdot|\pi^{**}, \mu) & , \pi = \pi^{**} \\ \sigma^0_{\mathscr{M}}(\cdot|\pi, \mu) & , \text{else} \end{cases} \quad \text{for all } \mu \in \Delta\Omega$$

$$\sigma'_{\mathscr{A}}(\cdot|\pi, m) = \begin{cases} \sigma_{\mathscr{A}}(\cdot|\pi^*, \tau(m)) & , m \in \text{supp}\left(\sigma'_{\mathscr{M}}\left(\pi^{**}, \mu_R(\cdot|\pi^*, \tau(m))\right)\right) \\ & \quad \text{and } \pi = \pi^{**} \\ \sigma^0_{\mathscr{A}}(\cdot|\pi, m) & , m \in \Delta\Omega \text{ and } \pi \neq \pi^{**} \end{cases}$$

$$\mu'_R(\cdot|\pi, m) = \begin{cases} \mu & , m \in \text{supp}\left(\sigma'_{\mathscr{M}}\left(\pi^{**}, \mu\right)\right) \text{ and } \pi = \pi^{**} \\ F & , m \in \Delta\Omega \text{ and } \pi \neq \pi^{**} \end{cases},$$

where $\left(\left(\sigma^0_I, \sigma^0_m\right), \sigma^0_a\right)$ and μ^0_R are defined as in the proof of Lemma 1.6, and where the mapping $\tau : \bigcup_{\mu \in \text{supp}(\pi^{**})} \text{supp}\left(\sigma'_{\mathscr{M}}\left(\pi^{**}, \mu\right)\right) \to \bigcup_{\mu \in \text{supp}(\pi^*)} \text{supp}\left(\sigma_{\mathscr{M}}\left(\pi^*, \mu\right)\right)$ is a bijective function defined as follows: If the sender takes message m after choosing π^* in the PBE E, the sender chooses message $\tau^{-1}(m)$ after taking π^{**} in the equilibrium candidate E' if the experiment's outcome is $\mu_R(\cdot|\pi^*, m)$.[36] The above profiles constitute a PBE E': First, the receiver's beliefs are updated according to Bayes' rule whenever possible. Moreover, since $\int u_R(a, \omega)\, d\mu'_R(\omega|\pi^{**}, m) = \int u_R(a, \omega)\, d\mu_R(\omega|\pi^*, \tau(m))$ for all $a \in A$ and all $m \in \text{supp}\left(\sigma'_{\mathscr{M}}\left(\pi^{**}, \mu_R(\cdot|\pi^*, \tau(m))\right)\right)$, and since $\sigma_{\mathscr{A}}$ is a best response given μ_R in any subgame after the sender chooses π^*, one can infer that $\sigma'_{\mathscr{A}}$ is optimal given μ'_R

36. Bijectivity of τ (or rather τ^{-1}) ensures that E' is fully revealing: If $\mu_R(\cdot|\pi^*, m) \neq \mu_R(\cdot|\pi^*, m')$ for some $m, m' \in \bigcup_{\mu \in \text{supp}(\pi^*)} \text{supp}\left(\sigma_{\mathscr{M}}\left(\pi^*, \mu\right)\right)$, but $\tau^{-1}(m) = \tau^{-1}(m')$, the sender would not fully reveal to the receiver whether he observed outcome $\mu_R(\cdot|\pi^*, m)$ or $\mu_R(\cdot|\pi^*, m')$ after choosing π^{**} in E'. Moreover, $\sigma'_{\mathscr{A}}(\cdot|\pi^{**}, m)$ is well-defined for all $m \in \bigcup_{\mu \in \text{supp}(\pi^{**})} \text{supp}\left(\sigma'_{\mathscr{M}}\left(\pi^{**}, \mu\right)\right)$ since the communication rule is fully revealing on the equilibrium path and the receiver updates her beliefs according to Bayes' rule: For any such m, which the sender sends after choosing π^{**} in the candidate for a fully revealing equilibrium E' , there exists a message $\tau(m)$ which the sender takes in the PBE E after choosing π^* so that the receiver's belief after observing π^* and $\tau(m)$ in E equals the outcome of the experiment π^{**} conditional on which the sender transmits m in E'.

in any subgame after he takes π^{**}. Besides, $\sigma'_{\mathscr{A}}$ is also a best response given μ'_R in any subgame after the sender chooses an experiment $\pi \neq \pi^{**}$ (it's the babbling outcome). Hence, $\sigma'_{\mathscr{A}}$ is optimal given μ'_R. Furthermore, one obtains that

$$\begin{aligned}
&\iint u_S(a,\omega)\ d\sigma'_{\mathscr{A}}\left(a|\pi^{**},m\right)\ d\mu(\omega) \\
&= \iint u_S(a,\omega)\ d\sigma_{\mathscr{A}}\left(a|\pi^{*},\tau(m)\right)\ d\mu(\omega) \\
&= \int \underbrace{\iint u_S(a,\omega)\ d\sigma_{\mathscr{A}}\left(a|\pi^{*},\tau(m)\right)\ d\mu(\omega)}_{\geq \iint u_S(a,\omega)\ d\sigma_{\mathscr{A}}\left(a|\pi^{*},\tau\left(m'\right)\right)\ d\mu(\omega)\ \text{ for all } \mu' \in \operatorname{supp}\left(\pi^{*}\right)} d\Pr\left(\mu'|\pi^{*},\tau(m),\sigma_{\mathscr{M}}\right) \\
&\geq \iint u_S(a,\omega)\ d\sigma_{\mathscr{A}}\left(a|\pi^{*},\tau\left(m'\right)\right)\ d\mu(\omega) \\
&= \iint u_S(a,\omega)\ d\sigma'_{\mathscr{A}}\left(a|\pi^{**},m'\right)\ d\mu(\omega)
\end{aligned}$$

for all $m \in \operatorname{supp}\left(\sigma'_{\mathscr{M}}\left(\pi^{**},\mu\right)\right)$, $m' \in \Delta\Omega$ and $\mu \in \operatorname{supp}(\pi^{**})$. The first equality holds true because $\sigma'_{\mathscr{A}}(\cdot|\pi^{**},m) = \sigma_{\mathscr{A}}(\cdot|\pi^{*},\tau(m))$ for every $m \in \bigcup_{\mu\in\operatorname{supp}(\pi^{**})} \operatorname{supp}\left(\sigma'_{\mathscr{M}}\left(\pi^{**},\mu\right)\right)$. The second equality is obtained by the law of iterated expectations, with $\Pr\left(\cdot|\pi^{*},\tau(m),\sigma_{\mathscr{M}}\right)$ being the distribution over outcomes after the sender chooses π^{*} and given he takes message $\tau(m)$ according to $\sigma_{\mathscr{M}}$. The weak inequality is due to the fact that $\tau(m)$ is optimal for the sender in the PBE E after taking π^{*} and learning its outcome μ so that $\tau(m) \in \operatorname{supp}\left(\sigma_{\mathscr{M}}\left(\pi^{*},\mu\right)\right)$. The last two equalities follow from the same arguments as for the first two equalities applied in reverse order. So $\sigma'_{\mathscr{M}}$ is a best response given $\sigma'_{\mathscr{A}}$ in any subgame after the sender chooses π^{**}. Also, $\sigma'_{\mathscr{M}}$ is a best response given $\sigma'_{\mathscr{A}}$ in any subgame after the sender chooses an $\pi \neq \pi^{**}$ (it's the babbling outcome). Hence, $\sigma'_{\mathscr{M}}$ is optimal given $\sigma'_{\mathscr{A}}$. By the law of iterated expectations, it holds for all $i \in \{S,R\}$ that

$$\begin{aligned}
&\iiiint u_i(a,\omega)\ d\sigma'_{\mathscr{A}}\left(a|\pi^{**},m\right)\ d\mu(\omega)\ d\sigma'_{\mathscr{M}}\left(m|\pi^{**},\mu\right)\ d\pi^{**}\left(\mu\right) \qquad (1.A.1)\\
&= \int\cdots\int u_i(a,\omega)\ d\sigma_{\mathscr{A}}\left(a|\pi^{*},\tau(m)\right)\ d\mu(\omega)\ d\Pr\left(\mu'|\pi^{*},\tau(m),\sigma_{\mathscr{M}}\right) \\
&\qquad\qquad d\Pr\left(\tau(m)|\pi^{*}\right) \\
&= \iiiint u_i(a,\omega)\ d\sigma_{\mathscr{A}}\left(a|\pi^{*},\tau(m)\right)\ d\mu(\omega)\ d\sigma_{\mathscr{M}}\left(\tau(m)|\pi^{*},\mu\right)\ d\pi^{*}\left(\mu\right),
\end{aligned}$$

so each agent's expected utilities after π^{**} in E' and after π^{*} in E coincide. This yields

$$\begin{aligned}
&\iiiint u_S(a,\omega)\ d\sigma'_{\mathscr{A}}\left(a|\pi^{**},m\right)\ d\mu(\omega)\ d\sigma'_{\mathscr{M}}\left(m|\pi^{**},\mu\right)\ d\pi^{**}\left(\mu\right)-c\left(\pi^{**}\right) \qquad \text{(1.A.2)}\\
&\geq \iiiint u_S(a,\omega)\ d\sigma_{\mathscr{A}}\left(a|\pi^{*},\tau(m)\right)\ d\mu(\omega)\ d\sigma_{\mathscr{M}}\left(\tau(m)|\pi^{*},\mu\right)\ d\pi^{*}\left(\mu\right)\\
&\qquad -c\left(\pi^{*}\right)\\
&\geq \mathscr{U}_S\left(E^0\right)\\
&= \iiiint u_S(a,\omega)\ d\sigma'_{\mathscr{A}}\left(a|\pi^{0},m\right)\ d\mu(\omega)\ d\sigma'_{\mathscr{M}}\left(m|\pi^{0},\mu\right)\ d\pi^{0}\left(\mu\right)-c\left(\pi^{0}\right)\\
&\geq \iiiint u_S(a,\omega)\ d\sigma'_{\mathscr{A}}\left(a|\pi,m\right)\ d\mu(\omega)\ d\sigma'_{\mathscr{M}}\left(m|\pi,\mu\right)\ d\pi\left(\mu\right)-c\left(\pi\right)
\end{aligned}$$

for all $\pi \in \Pi$. The first inequality follows from $c(\pi^{**}) \leq c(\pi^{*})$. Since E is a PBE where the sender chooses π^{*}, the term in the second line equals $\mathscr{U}_S(E)$. The second inequality immediately follows from the definition of E^0 (cf. the proof of Lemma 1.6). Since the sender's expected utility is constant across all subgames after the sender chooses some $\pi \neq \pi^{**}$ (as the babbling outcome is implemented in any such subgame), the last inequality results from the fact that $c(\pi) \geq c\left(\pi^0\right)$ by Assumption 1.2 as π^0 is a Blackwell garbling of π. As a consequence, the $\sigma'_{\mathscr{I}}$ is optimal given $\left(\sigma'_{\mathscr{M}}, \sigma'_{\mathscr{A}}\right)$. This concludes the proof that $\left\{\left(\left(\sigma'_{\mathscr{I}}, \sigma'_{\mathscr{M}}\right), \sigma'_{\mathscr{A}}\right), \mu'_R\right\}$ forms a PBE E', which is fully revealing and has a pure-strategy information rule. (1.A.1) implies that $\mathscr{U}_R\left(E'\right) = \mathscr{U}_R(E)$, and that $\mathscr{U}_S\left(E'\right) \geq \mathscr{U}_S(E)$ since $c(\pi^{**}) \leq c(\pi^{*})$. The PBE E is thus dominated by the fully revealing PBE E' with a pure-strategy information rule.

If c is strictly monotone, the first inequality in (1.A.2) is strict implying that any PBE with a pure-strategy information rule which is not fully revealing is strictly dominated by a fully revealing equilibrium.

Proof of Theorem 1.9. I apply Berge's maximum theorem to the optimization problem stated on page 15. Let $\mathscr{S}$ be the set of continuous mappings $\tilde{\sigma}_{\mathscr{A}} : \left(\Delta\Omega, d_{\text{LP}}\right) \to \left(\Delta A, d_{\text{LP}}\right)$, where d_{LP} is the Lévy–Prokhorov metric[37], endowed with the sup metric. Let $C : \mathbb{R} \rightrightarrows \Pi \times \mathscr{S}$ be the correspondence defined by $C\left(\mathscr{U}_R\right) \equiv \left\{\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) \in \Pi \times \mathscr{S} | \left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) \text{ satisfies } (1.1)-(1.4).\right\}$ for each $\mathscr{U}_R \in \mathbb{R}$. Recall that Π is compact. Moreover, the set of continuous functions $\mathscr{S}$ endowed with the sup metric constitutes a compact space. So by Tychonoff's theorem, $\Pi \times \mathscr{S}$ is compact. To apply the maximum theorem, I need the following auxiliary lemmata:

Lemma 1.47. *If $C\left(\mathscr{U}_R\right) \neq \emptyset$, then $C\left(\mathscr{U}_R\right)$ is compact.*

Proof. Suppose $C\left(\mathscr{U}_R\right) \neq \emptyset$. Let's show that $C\left(\mathscr{U}_R\right)$ is closed: Fix some sequence $\left(\pi_n, \tilde{\sigma}_{\mathscr{A},n}\right)_{n\in\mathbb{N}}$ in $C\left(\mathscr{U}_R\right)$ with limit $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$. Let $\tilde{a}_n\left(\mu\right)(\tilde{a}\left(\mu\right))$ denote the random

37. See Dudley (1989), p.394 for the definition.

variable with distribution $\tilde{\sigma}_{\mathscr{A},n}(\cdot|\mu)(\tilde{\sigma}_{\mathscr{A}}(\cdot|\mu))$. I verify in sequence that $(\pi, \tilde{\sigma}_{\mathscr{A}})$ satisfies (1.1)–(1.4):

(1). Suppose some $a' \in \text{supp}(\tilde{\sigma}_{\mathscr{A}}(\mu))$ with $a' \notin \arg\max_{a\in A} \int u_R(a,\omega)\, d\mu(\omega)$ exists. Notice that $\Pr(\tilde{a}(\mu) \in (a'-\epsilon, a'+\epsilon)) = \tilde{\sigma}_{\mathscr{A}}(a'+\epsilon|\mu) - \tilde{\sigma}_a(a'-\epsilon|\mu) > 0$ for any $\epsilon > 0$ as $a' \in \text{supp}(\tilde{\sigma}_{\mathscr{A}}(\mu))$. Since the metric space $(\triangle A, d_{\text{LP}})$ is endowed with the Lévy-Prokhorov metric, $(\tilde{\sigma}_{\mathscr{A},n}(\cdot|\mu))$ converges in distribution to $\tilde{\sigma}_{\mathscr{A}}(\cdot|\mu)$ (see **billingsley1999convergence**, p.72). Take any $\epsilon > 0$ with $\Pr(\tilde{a}(\mu) \in \{a'-\epsilon, a'+\epsilon\}) = 0$, and note that $(a'-\epsilon, a'+\epsilon)$ is a continuity set of $\tilde{a}(\mu)$. Let ϵ^* denote the set of such ϵ. As $(\tilde{\sigma}_{\mathscr{A},n}(\cdot|\mu))$ converges in distribution to $\tilde{\sigma}_{\mathscr{A}}(\cdot|\mu)$, one gets $\Pr(\tilde{\sigma}_{\mathscr{A},n}(\cdot|\mu) \in (a'-\epsilon, a'+\epsilon)) \to \Pr(\tilde{\sigma}_{\mathscr{A}}(\cdot|\mu) \in (a'-\epsilon, a'+\epsilon))$ as $n \to \infty$ for any $\epsilon \in \epsilon^*$. As $\Pr(\tilde{\sigma}_{\mathscr{A}}(\cdot|\mu) \in (a'-\epsilon, a'+\epsilon)) > 0$, for all $\epsilon \in \epsilon^*$, there is an $n(\epsilon) \in \mathbb{N}$ with $\Pr(\tilde{\sigma}_{a,n}(\cdot|\mu) \in (a'-\epsilon, a'+\epsilon)) > 0$. So there is some $a'(\epsilon) \in (a'-\epsilon, a'+\epsilon)$ so that $a'(\epsilon) \in \arg\max_{a\in A} \int u_R(a,\omega)\, d\mu(\omega)$ because $(\pi_{n(\epsilon)}, \tilde{\sigma}_{\mathscr{A},n(\epsilon)})$ satisfies (1.1). By construction, $a'(\epsilon) \to a'$ as $\epsilon \to 0$.[38] Since $\arg\max_{a\in A} \int u_R(a,\omega)\, d\mu(\omega)$ is closed, this implies that $a' \in \arg\max_{a\in A} \int u_R(a,\omega)\, d\mu(\omega)$.

(2). Suppose $\iint u_S(a,\omega)\, d\tilde{\sigma}_{\mathscr{A}}(a|\mu)\, d\mu(\omega) < \iint u_S(a,\omega)\, d\tilde{\sigma}_{\mathscr{A}}(a|\mu')\, d\mu(\omega)$ for some $\mu, \mu' \in \text{supp}(\pi)$. So for all $n \in \mathbb{N}$, $\mu \notin \text{supp}(\pi_n)$ or $\mu' \notin \text{supp}(\pi_n)$. Take a subsequence $(\pi_{n_k}, \tilde{\sigma}_{\mathscr{A},n_k})_{k\in\mathbb{N}}$ of $(\pi_n, \tilde{\sigma}_{\mathscr{A},n})_{n\in\mathbb{N}}$ so that $\mu \notin \text{supp}(\pi_{n_k})$ for any k.[39] By analogous arguments as in the proof of step (1), there exists a subsubsequence $(\pi_{n_{k_l}}, \tilde{\sigma}_{\mathscr{A},n_{k_l}})_{l\in\mathbb{N}}$ and a sequence $(\mu_l)_l$ $((\mu'_l)_l)$ with $\mu_l \in \text{supp}(\pi_{n_{k_l}})$ $(\mu_l \in \text{supp}(\pi_{n_{k_l}}))$ for all l converging to μ (μ'). Since $\iint u_S(a,\omega)\, d\tilde{\sigma}_{\mathscr{A}}(a|\mu)\, d\mu(\omega) - \iint u_S(a,\omega)\, d\tilde{\sigma}_{\mathscr{A}}(a|\mu')\, d\mu(\omega)$ is continuous in (μ, μ') and $\iint u_S(a,\omega)\, d\tilde{\sigma}_{\mathscr{A}}(a|\mu)\, d\mu_l(\omega) - \iint u_S(a,\omega)\, d\tilde{\sigma}_{\mathscr{A}}(a|\mu')\, d\mu_l(\omega) \geq 0$ for all l because the sequence $(\pi_{n_{k_l}}, \tilde{\sigma}_{\mathscr{A},n_{k_l}})$ satisfies (1.2), one obtains that $\iint u_S(a,\omega)\, d\tilde{\sigma}_{\mathscr{A}}(a|\mu)\, d\mu(\omega) - \iint u_S(a,\omega)\, d\tilde{\sigma}_{\mathscr{A}}(a|\mu')\, d\mu(\omega) \geq 0$.

(3). As $\tilde{\mathscr{U}}_S$ is continuous, $\tilde{\mathscr{U}}_S(\pi_n, \tilde{\sigma}_{\mathscr{A},n}) \geq \mathscr{U}_S^0$ for all n implies that $\tilde{\mathscr{U}}_S(\pi, \tilde{\sigma}_{\mathscr{A}}) \geq \mathscr{U}_S^0$.

(4). By continuity of $\tilde{\mathscr{U}}_R$, $\tilde{\mathscr{U}}_R(\pi_n, \tilde{\sigma}_{\mathscr{A},n}) \geq \mathscr{U}_R$ for all n yields $\tilde{\mathscr{U}}_R(\pi, \tilde{\sigma}_{\mathscr{A}}) \geq \mathscr{U}_R$. Hence, $C(\mathscr{U}_R)$ is a closed subset of the compact set $\Pi \times \mathscr{S}$, so $C(\mathscr{U}_R)$ is compact. □

Lemma 1.48. *There is some $\mathscr{U}_R^{max} \in [\mathscr{U}_R^0, \infty)$ so that $C(\bar{\mathscr{U}}_R) \neq \emptyset$ iff $\bar{\mathscr{U}}_R \leq \mathscr{U}_R^{max}$.*

Proof. Note that $C(\mathscr{U}_R(E^0)) \neq \emptyset$ since $(\pi^0, \sigma^0_{\mathscr{A}}(\pi^0, \cdot)) \in C(\mathscr{U}_R(E^0))$, where $\sigma^0_{\mathscr{A}}$ is defined as in the proof of Lemma 1.6. Recall that E^0 is a fully revealing PBE with a pure-strategy information rule, so $(\pi^0, \sigma^0_{\mathscr{A}}(\pi^0, \cdot))$ satisfies (1.1) and (1.2). Also, $\tilde{\mathscr{U}}_S(\pi^0, \sigma^0_{\mathscr{A}}(\pi^0, \cdot)) = \mathscr{U}_S(E^0) \geq \mathscr{U}_S^0$ and

38. There exists a sequence in ϵ^* converging to zero because $\tilde{a}(\mu)$ has at most countably many realizations that occur with positive probability.

39. If such a subsequence does not exist, take any infinite subsequence $(\pi_{n_k}, \tilde{\sigma}_{\mathscr{A},n_k})_{k\in\mathbb{N}}$ of $(\pi_n, \tilde{\sigma}_{\mathscr{A},n})_{n\in\mathbb{N}}$ such that $\mu' \notin \text{supp}(\pi_{n_k})$ for any k, which must then exist, and switch the labels of μ and μ'.

$\tilde{\mathscr{U}}_R\left(\pi^0, \sigma^0_{\mathscr{A}}\left(\pi^0, \cdot\right)\right) = \mathscr{U}_R\left(E^0\right) = \mathscr{U}^0_R$, that is, $\left(\pi^0, \sigma_{\mathscr{A}}\left(\pi^0, \cdot\right)\right)$ fulfills (1.3) and (1.4). Second, $C\left(\mathscr{U}_R\right) \neq \emptyset$ implies $C\left(\mathscr{U}'_R\right) \neq \emptyset$ for all $\mathscr{U}'_R \leq \mathscr{U}_R$: If $\left(\pi^0, \sigma_{\mathscr{A}}\left(\pi^0, \cdot\right)\right) \in C\left(\mathscr{U}_R\right)$, then $\tilde{\mathscr{U}}_R\left(\pi^0, \sigma_{\mathscr{A}}\left(\pi^0, \cdot\right)\right) \geq \mathscr{U}_R \geq \mathscr{U}'_R$. Since (1.1)–(1.3) do not depend on $\mathscr{U}_R$, this yields $\left(\pi^0, \sigma_{\mathscr{A}}\left(\pi^0, \cdot\right)\right) \in C\left(\mathscr{U}'_R\right)$. Third, continuity of u_R and compactness of $A \times \Omega$ yield $\tilde{\mathscr{U}}_R\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) \leq \max_{(a,\omega)\in A\times\Omega} u_R(a, \omega) < \infty$ for any $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) \in \Pi \times \mathscr{S}$. Hence, there exists some $\mathscr{U}^{\max}_R \in \left[\mathscr{U}^0_R, \infty\right)$ such that $C\left(\mathscr{U}_R\right) \neq \emptyset$ if $\mathscr{U}_R < \mathscr{U}^{\max}_R$ and $C\left(\mathscr{U}_R\right) = \emptyset$ if $\mathscr{U}_R > \mathscr{U}^{\max}_R$. Finally, let's verify that $C\left(\mathscr{U}^{\max}_R\right) \neq \emptyset$: Consider an increasing sequence $\left(\mathscr{U}_{R,n}\right)_{n\in\mathbb{N}}$ with limit $\mathscr{U}^{\max}_R$, and a sequence $\left(\pi_n, \tilde{\sigma}_{\mathscr{A},n}\right)_{n\in\mathbb{N}}$ satisfying $\tilde{\mathscr{U}}_R\left(\pi_n, \tilde{\sigma}_{\mathscr{A},n}\right) = \mathscr{U}_{R,n}$ and (1.1)–(1.3) for all n. Such a sequence exists since $C\left(\mathscr{U}_R\right) \neq \emptyset$ for all $\mathscr{U}_R < \mathscr{U}^{\max}_R$. By compactness of $\Pi \times \mathscr{S}$, there exists a convergent subsequence of $\left(\pi_n, \tilde{\sigma}_{\mathscr{A},n}\right)_{n\in\mathbb{N}}$ with limit $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$. By continuity of $\tilde{\mathscr{U}}_R$, it follows that $\tilde{\mathscr{U}}_R\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) = \mathscr{U}^{\max}_R$. From the proof of Lemma 1.47, $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$ satisfies (1.1)–(1.3) so that $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) \in C\left(\mathscr{U}^{\max}_R\right)$. □

Lemma 1.49. *C is continuous in $\mathscr{U}_R$ on $\left(-\infty, \mathscr{U}^{max}_R\right]$.*

Proof. I show that C is both upper and lower hemicontinuous at any $\mathscr{U}_R \in \left(-\infty, \mathscr{U}^{\max}_R\right]$: Take any sequence $\left(\mathscr{U}_{R,n}\right)_n$ converging to $\mathscr{U}_R$, any tuple $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) \in \Pi \times \mathscr{S}$, and any sequence $\left(\pi_n, \tilde{\sigma}_{\mathscr{A},n}\right)_n$ so that $\left(\pi_n, \tilde{\sigma}_{\mathscr{A},n}\right) \in C\left(\mathscr{U}_{R,n}\right)$ for all n converging to $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$. Suppose $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) \notin C\left(\mathscr{U}_{R,n}\right)$. By Lemma 1.47, since $\left(\pi_n, \tilde{\sigma}_{\mathscr{A},n}\right)_n$ satisfies (1.1)–(1.3), so does its limit. So $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) \notin C\left(\mathscr{U}_R\right)$ means that $\tilde{\mathscr{U}}_R\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) < \mathscr{U}_R$. Continuity of $\tilde{\mathscr{U}}_R$ yields $\tilde{\mathscr{U}}_R\left(\pi_n, \tilde{\sigma}_{\mathscr{A},n}\right) < \mathscr{U}_{R,n}$ for large n—a contradiction. This proves upper hemicontinuity. Take again a sequence $\left(\mathscr{U}_{R,n}\right)_n$ converging to $\mathscr{U}_R$ and any $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) \in C\left(\mathscr{U}_R\right)$. If $\tilde{\mathscr{U}}_R\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) > \mathscr{U}_R$, then it follows that $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) \in C\left(\mathscr{U}_{R,n}\right)$ for large n. So there is a subsequence $\left(\mathscr{U}_{R,n_l}\right)_l$ and a sequence $\left(\pi_l, \tilde{\sigma}_{\mathscr{A},l}\right)_l$ with $\left(\pi_l, \tilde{\sigma}_{\mathscr{A},l}\right) = \left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$ and $\tilde{\mathscr{U}}_R\left(\pi_l, \tilde{\sigma}_{\mathscr{A},l}\right) > \left(\mathscr{U}_{R,l}\right)$ for all l, which converges to $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$. If $\tilde{\mathscr{U}}_R\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) = \mathscr{U}_R$, there is, by continuity of $\tilde{\mathscr{U}}_R$, some $\left(\pi_n, \tilde{\sigma}_{\mathscr{A},n}\right)_n$ satisfying (1.1)–(1.3) and $\tilde{\mathscr{U}}_R\left(\pi_n, \tilde{\sigma}_{\mathscr{A},n}\right) = \mathscr{U}_{R,n}$ for all n with limit $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$. This proves lower hemicontinuity. □

So the objective function $\tilde{\mathscr{U}}_S$ is continuous in $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$, and the correspondence C is compact-valued, non-empty and continuous on $\left(-\infty, \mathscr{U}^{\max}_R\right]$. It follows by Berge's maximum theorem that $\arg\max_{(\pi,\tilde{\sigma}_{\mathscr{A}})\in C(\mathscr{U}_R)} \tilde{\mathscr{U}}_S\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$ is non-empty for every $\mathscr{U}_R \in \left(-\infty, \mathscr{U}^{\max}_R\right]$, and $\mathscr{U}^*_S\left(\bar{\mathscr{U}}_R\right) \equiv \max_{(\pi,\tilde{\sigma}_{\mathscr{A}})\in C(\mathscr{U}_R)} \tilde{\mathscr{U}}_S\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$ is continuous in $\mathscr{U}_R$ on $\left(-\infty, \mathscr{U}^{\max}_R\right]$. This proves the existence part of the theorem. To prove the compactness part, notice first that the set of Pareto efficient implementable payoff profiles is $\bigcup_{\mathscr{U}_R\in(-\infty,\mathscr{U}^{\max}_R]} \mathscr{U}^*_S\left(\mathscr{U}_R\right)$. Moreover, observe that $\min_{(a,\omega)\in A\times\Omega} u_S(a, \omega) \leq \mathscr{U}^*_S\left(\mathscr{U}_R\right) \leq \max_{(a,\omega)\in A\times\Omega} u_S(a, \omega)$ for all $\mathscr{U}_R \in \left(-\infty, \mathscr{U}^{\max}_R\right]$, where the min- and max-function are well-defined by continuity of u_S and compactness of $A \times \Omega$. As a result, continuity of $\mathscr{U}^*_S$ implies that $\bigcup_{\mathscr{U}_R\in(-\infty,\mathscr{U}^{\max}_R]} \mathscr{U}^*_S\left(\mathscr{U}_R\right)$ is compact.

Proof of Lemma 1.11. Let $K = \left\{V \in \mathscr{P}_S : \mathscr{U}_S^{\min}(V) \geq V' \text{ for all } V' \in \mathscr{P}_S : V' \leq V\right\}$, and note that $K \subseteq \mathscr{P}_S \subset \mathscr{R}$. Since $\mathscr{P}_S$ is compact by Theorem 1.9 and thus bounded, $\sup K$ exists by the Supremum Property. Moreover, $\sup K \in \mathscr{P}_S$. Consequently, $\left(\mathscr{U}_S, \mathscr{U}_R\right) \in \mathscr{P}$ immediately follows from $\mathscr{U}_S \in \mathscr{P}_S$ by (1.5). Uniqueness is obvious (as suprema are unique).

1.A.3 Section 1.3.3

Proof of Theorem 1.13. Analogously to Theorem 1.9, the proof proceeds by applying Berge's maximum theorem to the optimization problem on page 17. Details are therefore omitted.

Proof of Theorem 1.14. Let $\left(\mathscr{U}_S^*, \mathscr{U}_R^*\right)$ be the unique Pareto efficient feasible payoff profile. By feasibility, there is some $\left(\bar{\sigma}_{\mathscr{I}}, \bar{\sigma}_{\mathscr{A}}\right)$ so that $\mathscr{U}_i^* = \bar{\mathscr{U}}_i\left(\bar{\sigma}_{\mathscr{I}}, \bar{\sigma}_{\mathscr{A}}\right)$ for all i. First, the strategy profile $\left(\left(\bar{\sigma}_I, \sigma_m^{\mathrm{FR}}\right), \bar{\sigma}_{\mathscr{A}}\right)$ together with some beliefs μ_R form a PBE E: Suppose the receiver has an incentive to deviate. Then, there is some $\sigma_{\mathscr{A}}$ with $\mathscr{U}_R\left(\left\{\left(\left(\bar{\sigma}_{\mathscr{I}}, \sigma_{\mathscr{M}}^{\mathrm{FR}}\right), \sigma_{\mathscr{A}}\right), \mu_R\right\}\right) > \mathscr{U}_R^*$. As $\sigma_{\mathscr{M}}^{\mathrm{FR}}$ is fully revealing, $\mathscr{U}_i\left(\left\{\left(\left(\bar{\sigma}_{\mathscr{I}}, \sigma_{\mathscr{M}}^{\mathrm{FR}}\right), \sigma_{\mathscr{A}}\right), \mu_R\right\}\right) = \bar{\mathscr{U}}_i\left(\bar{\sigma}_{\mathscr{I}}, \sigma_{\mathscr{A}}\right)$ for all i, i.e., the payoff profile induced by the receiver's deviation is feasible. This yields $\bar{\mathscr{U}}_S\left(\bar{\sigma}_{\mathscr{I}}, \sigma_{\mathscr{A}}\right) < \mathscr{U}_S^*$ as $\left(\bar{\mathscr{U}}_S\left(\bar{\sigma}_{\mathscr{I}}, \sigma_{\mathscr{A}}\right), \bar{\mathscr{U}}_R\left(\bar{\sigma}_{\mathscr{I}}, \sigma_{\mathscr{A}}\right)\right)$ is feasible and would otherwise strictly dominate $\left(\mathscr{U}_S^*, \mathscr{U}_R^*\right)$—a contradiction since $\left(\mathscr{U}_S^*, \mathscr{U}_R^*\right)$ is Pareto efficient feasible. By compactness of the set of Pareto efficient feasible payoff profiles (cf. Theorem 1.13), there exists thus a Pareto efficient feasible payoff profile $\left(\mathscr{U}_S', \mathscr{U}_R'\right)$ with $\mathscr{U}_R' \geq \bar{\mathscr{U}}_R\left(\bar{\sigma}_{\mathscr{I}}, \sigma_{\mathscr{A}}\right) > \mathscr{U}_R^*$ and $\mathscr{U}_S' \leq \bar{\mathscr{U}}_S\left(\bar{\sigma}_{\mathscr{I}}, \sigma_{\mathscr{A}}\right) < \mathscr{U}_S^*$, contradicting uniqueness of the Pareto efficient feasible payoff profile. Now, suppose the sender has a profitable deviation $\left(\sigma_{\mathscr{I}} \sigma_{\mathscr{M}}\right)$, that is, $\mathscr{U}_S\left(\left\{\left(\left(\sigma_{\mathscr{I}}, \sigma_{\mathscr{M}}\right), \bar{\sigma}_{\mathscr{A}}\right), \mu_R\right\}\right) > \mathscr{U}_S^*$. Notice that there is some $\hat{\sigma}_{\mathscr{A}}$ so that $\mathscr{U}_i\left(\left\{\left(\left(\sigma_{\mathscr{I}}, \sigma_{\mathscr{M}}\right), \bar{\sigma}_{\mathscr{A}}\right), \mu_R\right\}\right) = \bar{\mathscr{U}}_i\left(\sigma_{\mathscr{I}}, \hat{\sigma}_{\mathscr{A}}\right)$ for all i, i.e., the payoff profile induced by the sender's deviation is feasible. One can conclude from this that either $\left(\mathscr{U}_S^*, \mathscr{U}_R^*\right)$ is not Pareto efficient feasible, or there exists another Pareto efficient feasible payoff profile, contradicting uniqueness. By construction, it holds that $\mathscr{U}_i(E) = \mathscr{U}_i^*$ for all i. Hence, $\left(\mathscr{U}_S^*, \mathscr{U}_R^*\right)$ is implementable. Since $\left(\mathscr{U}_S^*, \mathscr{U}_R^*\right)$ is also Pareto efficient feasible, it is Pareto efficient implementable. Finally, suppose $\left(\mathscr{U}_S^*, \mathscr{U}_R^*\right)$ is not the unique Pareto efficient implementable payoff profile, i.e., there is another Pareto efficient implementable payoff profile $\left(\mathscr{U}_S'', \mathscr{U}_R''\right)$. But then, since $\left(\mathscr{U}_S^*, \mathscr{U}_R^*\right)$ is also feasible, there exists another Pareto efficient feasible payoff profile by Theorem 1.13—a contradiction.

Proof of Theorem 1.15. Fix some Pareto efficient feasible payoff profile $\left(\mathscr{U}_S^*, \mathscr{U}_R^*\right)$ and suppose that the conditions stated in Theorem 1.15 hold true. In particular, $\mathrm{supp}(\bar{\sigma}_{\mathscr{I}}) = \left\{\pi^{\mathrm{full}}\right\}$ implies that the best feasible payoff profile for agent i is generated if the social planner fully learns the state and chooses the optimal action for agent i per state, i.e., this payoff profile is $\left(\int u_S\left(a_i^*(\omega), \omega\right) d\mu_0(\omega) - c\left(\pi^{\mathrm{full}}\right), \int u_R\left(a_i^*(\omega), \omega\right) d\mu_0(\omega)\right)$. Since there are at least two Pareto efficient feasible payoff profiles, the sender-optimal and receiver-optimal ones do not coincide so that the set of states with $a_S^*(\omega) \neq a_R^*(\omega)$ is of positive measure. If $\left(\mathscr{U}_S^*, \mathscr{U}_R^*\right)$ is implementable, there is a PBE $\left\{\left(\left(\bar{\sigma}_{\mathscr{I}}, \sigma_{\mathscr{M}}^{\text{FR}}\right), \bar{\sigma}_{\mathscr{A}}\right), \mu_R\right\}$ for some such $\left(\bar{\sigma}_{\mathscr{I}}, \bar{\sigma}_{\mathscr{A}}\right)$. However, if $\left(\mathscr{U}_S^*, \mathscr{U}_R^*\right)$ is the receiver-optimal feasible payoff profile, the sender has an incentive to misreport in the communication stage to generate the sender-optimal outcome. Similarly, if $\left(\mathscr{U}_S^*, \mathscr{U}_R^*\right)$ is any other Pareto efficient feasible payoff profile, the receiver has an incentive to deviate in the action stage to generate the receiver-optimal outcome.

Proof of Corollary 1.16. The "if" direction holds due to Theorem 1.14. The "only if" direction follows from Theorem 1.15: Fix a Pareto efficient feasible payoff profile $\left(\mathscr{U}_S^*, \mathscr{U}_R^*\right)$. From the optimization problem on page 17, it follows that there is some $\left(\bar{\sigma}_{\mathscr{I}}^*, \bar{\sigma}_{\mathscr{A}}^*\right)$ solving $\max_{\left(\bar{\sigma}_{\mathscr{I}}, \bar{\sigma}_{\mathscr{A}}\right)} \alpha \bar{\mathscr{U}}_S\left(\bar{\sigma}_{\mathscr{I}}, \bar{\sigma}_{\mathscr{A}}\right) + (1-\alpha)\bar{\mathscr{U}}_R\left(\bar{\sigma}_{\mathscr{I}}, \bar{\sigma}_{\mathscr{A}}\right)$, with $\alpha \in [0,1]$, so that $\mathscr{U}_i^* = \bar{\mathscr{U}}_i\left(\bar{\sigma}_{\mathscr{I}}^*, \bar{\sigma}_{\mathscr{A}}^*\right)$ for all i. Since $c = 0$, one obtains that

$$\begin{aligned}
&\alpha \bar{\mathscr{U}}_S\left(\bar{\sigma}_{\mathscr{I}}, \bar{\sigma}_{\mathscr{A}}\right) + (1-\alpha)\bar{\mathscr{U}}_R\left(\bar{\sigma}_{\mathscr{I}}, \bar{\sigma}_{\mathscr{A}}\right) \\
&\quad = \iiiint \alpha u_S(a,\omega) + (1-\alpha)u_R(a,\omega)\ d\bar{\sigma}_{\mathscr{A}}(a|\pi,\mu)\ d\mu(\omega)\ d\pi\left(\mu\right)\ d\bar{\sigma}_{\mathscr{I}}(\pi).
\end{aligned}$$

Since cost are zero and $A^*(\omega, \alpha u_S + (1-\alpha)u_R) \cap A^*(\omega', \alpha u_S + (1-\alpha)u_R) = \emptyset$ for all $\omega \neq \omega'$, the objective function is maximized if and only if $\text{supp}(\bar{\sigma}_{\mathscr{I}}) = \{\pi^{\text{full}}\}$.

Appendix 1.B Proofs: Optimal Experiments

1.B.1 Section 1.4.1

Proof of Theorem 1.19. Suppose the sender chooses an experiment π not satisfying probabilistic independence in any fully revealing PBE E with a pure-strategy information rule generating Pareto efficient implementable payoffs $(\mathscr{U}_S, \mathscr{U}_R)$. Let μ^* be the set of outcomes $\mu \in \text{supp}(\pi)$ with $\mu = \int \nu\, d\lambda_\mu(\nu)$ for some λ_μ not being equal to δ_μ a.e., and let λ_{μ^*} be the conditional distribution over outcomes of π given $\mu \in \mu^*$. Since μ^* is of positive measure $p \in (0,1]$, there exists some $\lambda' \in \triangle\left(\triangle\Omega\right)$ with $\pi = p\lambda_{\mu^*} + (1-p)\lambda'$. Let $\pi' \equiv p\int_{\mu^*} \lambda_\mu\, d\lambda_{\mu^*}\left(\mu\right) + (1-p)\lambda'$ be the experiment that results from replacing mass p of outcomes in μ^* by the corresponding distribution over outcomes λ_μ. By construction, π' satisfies probabilistic independence. Furthermore, π is a Blackwell garbling of π'. Implementability of $\left(\mathscr{U}_S, \mathscr{U}_R\right)$ yields existence of some $\tilde{\sigma}_{\mathscr{A}}$ so that $\mathscr{U}_i = \tilde{\mathscr{U}}_i\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$ for all i and implementability of $\left(\tilde{\mathscr{U}}_S\left(\pi', \tilde{\sigma}_{\mathscr{A}}\right), \tilde{\mathscr{U}}_R\left(\pi', \tilde{\sigma}_{\mathscr{A}}\right)\right)$ because $\text{supp}(\pi') \subseteq \text{supp}(\pi)$. Incentive compatibility (1.2) of $\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$ implies $\tilde{\mathscr{U}}_S\left(\pi', \tilde{\sigma}_{\mathscr{A}}\right) = \tilde{\mathscr{U}}_S\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$ since cost are zero. Since π is a Blackwell garbling of π', one gets $\tilde{\mathscr{U}}_R\left(\pi', \tilde{\sigma}_{\mathscr{A}}\right) \geq \tilde{\mathscr{U}}_R\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$. So $\left(\mathscr{U}_S, \mathscr{U}_R\right)$ is either dominated by $(\tilde{\mathscr{U}}_S\left(\pi', \tilde{\sigma}_{\mathscr{A}}\right), \tilde{\mathscr{U}}_R\left(\pi', \tilde{\sigma}_{\mathscr{A}}\right))$, or it is generated by a fully revealing PBE with a pure-strategy information rule in which the sender chooses the experiment π'.

Proof of Corollary 1.20. If $\text{supp}(\mu_0) = \{\omega_1, \omega_2\}$, any experiment π satisfying probabilistic independence can have most two different outcomes: Suppose not, i.e., there are at least three different outcomes $\mu_1, \mu_2, \mu_3 \in \text{supp}(\pi)$. Define $w_i \equiv \left(\mu_i\left(\omega_1\right), \mu_i\left(\omega_2\right)\right)'$ for each i, and notice that $\mu_i(\omega) = 0$ for all $\omega \notin \text{supp}(\mu_0)$. The 2-dimensional vectors w_1, w_2 and w_3 are linearly independent implying existence of some $\beta_1, \beta_2, \beta_3 \in \mathbb{R}$, not all zero, so that $\sum_{i=1}^{3} \beta_i \cdot w_i = 0_2$. Without loss of generality, suppose $\beta_1 < 0$ and $\beta_2, \beta_3 \geq 0$. Rearranging $\sum_{i=1}^{3} \beta_i \cdot w_i = 0_2$ yields $w_1 = \sum_{i=2}^{3} \frac{\beta_i}{-\beta_1} \cdot w_i$, where $\frac{\beta_i}{-\beta_1} \geq 0$ for $i = 2, 3$ and $\sum_{i=2}^{3} \frac{\beta_i}{-\beta_1} = 1$ as $\sum_{i=1}^{3} \beta_i = 0$. Since this holds true for any three outcomes, there is a positive measure of outcomes in $\text{supp}(\pi)$ that can be represented as a convex combination of other outcomes in $\text{supp}(\pi)$—a contradiction. If cost are zero, an immediate consequence of Theorem 1.19 is that any Pareto efficient implementable payoff profile is generated in a fully revealing PBE in which the sender chooses some π with $\text{supp}(\pi) \leq 2$.

1.B.2 Section 1.4.2

Proof of Lemma 1.24. If π is no extreme point, there exist some $\alpha \in [0, 1]$, and some $\pi_1, \pi_2 \in \Pi$ with $\pi_1 \neq \pi_2$ such that $\pi = \alpha\pi_1 + (1-\alpha)\pi_2$. But then, $\int v d\pi_1(v) = \int v d\pi_2(v)$, so π does not satisfy strong probabilistic independence.

If π does not satisfy strong probabilistic independence, there exist some $p \in (0, 1)$, and some $\lambda_1, \lambda_2, \lambda_1', \lambda_2' \in \triangle\left(\triangle\Omega\right)$ with $p\lambda_i + (1-p)\lambda_i' = \pi$ for all i such that λ_1 and λ_2 are not equal a.e., but $\int_\Pi v\, d\lambda_1(v) = \int_\Pi v\, d\lambda_2(v)$. Let $\pi_1 \equiv p\lambda_2 + (1-p)\lambda_1'$ be the experiment resulting from π by replacing mass p of the distribution over outcomes λ_1 by mass p of λ_2, and $\pi_2 \equiv p\lambda_1 + (1-p)\lambda_2'$. Note that $\frac{\pi_1+\pi_2}{2} = \frac{p\lambda_1+(1-p)\lambda_1'+\lambda_2+(1-p)\lambda_2'}{2} = \pi$, so π is no extreme point of Π.

Proof of Theorem 1.25. Suppose the sender chooses an experiment π not being a convex combination of two experiments satisfying strong probabilistic independence in any fully revealing PBE E with a pure-strategy information rule generating Pareto efficient implementable payoffs $(\mathscr{U}_S, \mathscr{U}_R)$. In particular, π does not satisfy strong probabilistic independence as $\pi = \alpha \cdot \pi + (1-\alpha) \cdot \pi$ for any $\alpha \in [0, 1]$. By Lemma 1.24, π is no extreme point, so there exist π_1, π_2 with $\pi = \frac{\pi_1+\pi_2}{2}$. By assumption, π_1 or π_2 does not satisfy strong probabilistic independence, so for concreteness, suppose π_2 does so. By Lemma 1.24, π_2 is no extreme point, so there exist π_3, π_4 with $\pi_2 = \frac{\pi_3+\pi_4}{2}$. Thus,$\pi = \frac{2\pi_1+\pi_3+\pi_4}{4}$. Let Π^{conv} be the set of all convex combinations $\sum_{i\in\{1,3,4\}} \alpha_i \pi_i$. Implementability of $\left(\mathscr{U}_S, \mathscr{U}_R\right)$ implies existence of some $\tilde{\sigma}_{\mathscr{A}}$ so that $\mathscr{U}_i = \tilde{\mathscr{U}}_i\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right)$ for all i. Also, it implies implementability of $\left(\tilde{\mathscr{U}}_S\left(\pi', \tilde{\sigma}_{\mathscr{A}}\right), \tilde{\mathscr{U}}_R\left(\pi', \tilde{\sigma}_{\mathscr{A}}\right)\right)$ as $\text{supp}(\pi') \subseteq \text{supp}(\pi)$ ensures that (1.2) holds true for any $\pi' \in \Pi^{\text{conv}}$. Denote by $\Pi\left(\mathscr{U}_R\right)$ the set of experiments $\pi' \in \Pi^{\text{conv}}$ with $\tilde{\mathscr{U}}_R\left(\pi', \tilde{\sigma}_{\mathscr{A}}\right) = \mathscr{U}_R$. By linearity of $\tilde{\mathscr{U}}_R\left(\cdot, \tilde{\sigma}_{\mathscr{A}}\right)$ on Π^{conv}, this set forms a straight line in the 2-simplex Π^{conv}. Linearity of $\tilde{\mathscr{U}}_S\left(\cdot,\tilde{\sigma}_{\mathscr{A}}\right)$ on Π^{conv} implies that one of the endpoints of this line π'' maximizes $\tilde{\mathscr{U}}_S\left(\cdot,\tilde{\sigma}_{\mathscr{A}}\right)$ on Π^{conv} on that line. Moreover, $\alpha_1\alpha_3\alpha_4 = 0$ at any such endpoint, that is, π'' a convex combination of π_1 and π_3, π_1 and π_4, or π_3 and π_4. By construction, $\tilde{\mathscr{U}}_S\left(\pi'',\tilde{\sigma}_{\mathscr{A}}\right) \geq \mathscr{U}_S$ as $\pi'' \in \Pi\left(\mathscr{U}_R\right)$. If $\tilde{\mathscr{U}}_S\left(\pi'',\tilde{\sigma}_{\mathscr{A}}\right) > \mathscr{U}_S$, $\left(\mathscr{U}_S,\mathscr{U}_R\right)$ is dominated by the implementable payoff profile $\left(\tilde{\mathscr{U}}_S\left(\pi'',\tilde{\sigma}_{\mathscr{A}}\right),\mathscr{U}_R\right)$ contradicting Pareto efficient implementability of the former one. If $\tilde{\mathscr{U}}_S\left(\pi'',\tilde{\sigma}_{\mathscr{A}}\right) = \mathscr{U}_S$, then $\left(\mathscr{U}_S,\mathscr{U}_R\right)$ is generated by a fully revealing PBE with a pure-strategy information rule in which the sender chooses the experiment π'' being a convex combination of two experiments. If these two experiments satisfy strong probabilistic independence, the proof is complete. If not, one can repeat the above argumentation taking $\pi^1 = \pi''$ as the new initial experiment instead of π. One can rerun this procedure until either $\left(\mathscr{U}_S,\mathscr{U}_R\right)$ is shown to be dominated or until one finds an experiment π''^n satisfying the properties mentioned in the theorem's statement. This process ends eventually: If $\left(\mathscr{U}_S,\mathscr{U}_R\right)$ is not shown to be dominated in any (finite) round n and the constructed experiment π''^n does not satisfy (i), then $\lim_{n\to\infty}\pi''^n$ is an extreme point of Π^{conv} thus also of $\triangle\left(\triangle\Omega\right)$, hence satisfying strong probabilistic independence. Consequently, $\left(\mathscr{U}_S,\mathscr{U}_R\right)$ is generated by a fully revealing PBE with a pure-strategy information rule where the sender takes $\lim_{n\to\infty}\pi''^n$.

Proof of Corollary 1.26. Suppose $\text{supp}(\mu_0) = \{\omega_1,\dots,\omega_n\}$ for some $n \in \mathbb{N}$. Verify that $|\text{supp}(\pi)| \leq \text{supp}(\mu_0) = n$ for all π satisfying strong probabilistic independence. If not, there exists some π satisfying strong probabilistic independence and $|\text{supp}(\pi)| \geq n+1$. There are at least $n+1$ different distributions over outcomes $\lambda_1,\dots,\lambda_{n+1} \in \triangle\left(\triangle\Omega\right)$ of positive measure: $\sum_{i=1}^{n+1}\alpha_i\lambda_i + \left(1-\sum_{i=1}^{n+1}\alpha_i\right)\lambda' = \pi$ for some $\alpha_1,\dots,\alpha_{n+1} \geq 0$ with $\sum_{i=1}^{n+1}\alpha_i \leq 1$. Define the vector $v_i \equiv \left(\int_\Pi \mu\left(\omega_1\right)\,d\lambda_i\left(\mu\right),\dots,\int_\Pi \mu\left(\omega_n\right)\,d\lambda_i\left(\mu\right)\right)'$ for all $i \in \{1,\dots,n+1\}$. As $\int_\Pi \mu\left(\omega\right)\,d\lambda_i\left(\mu\right) = 0$ for all $\omega \notin \Omega$, the posterior distribution of the state given the distribution over outcomes λ_i is fully characterized by v_i. Since $n+1$ n-dimensional vectors are linearly dependent, there are numbers $\beta_1,\dots,\beta_{n+1} \in \mathbb{R}$, not all zero, so that $\sum_{i=1}^{n+1}\beta_i\cdot v_i = 0_n$. Note that $\sum_{i=1}^{n+1}\beta_i = 0$ as $\sum_{j=1}^{n}\int_\Pi \mu\left(\omega_j\right)\,d\lambda_i\left(\mu\right) = 1$ for all i. Particularly, there is some i with $\beta_i < 0$. Rearranging $\sum_{i=1}^{n+1}\beta_i\cdot v_i = 0_n$ yields $\sum_{i:\beta_i<0}\frac{\beta_i}{-\sum_{j:\beta_j<0}\beta_j}\cdot v_i = \sum_{i:\beta_i\geq 0}\frac{\beta_i}{\sum_{j:\beta_j<0}\beta_j}\cdot v_i$, where $\frac{\beta_i}{-\sum_{j:\beta_j<0}\beta_j} \in [0,1]$ for all $\beta_i < 0$ and $\frac{\beta_i}{\sum_{j:\beta_j<0}\beta_j} \in [0,1]$ for all $\beta_i \geq 0$. Hence, there are two convex combinations (over convex combinations) over outcomes which can be represented by one another, that is, π does not satisfy strong probabilistic independence. With that, it follows from Theorem 1.25 that any Pareto efficient implementable payoff profile is generated by a fully revealing PBE in which the sender chooses a experiment with at most $n+n = 2n$ different outcomes.

Appendix 1.C Proofs: Optimality of Bi-Pooling Policies

Proof of Lemma 1.31. Denote $\bar{\omega}(\mu) = \int \omega \, d\mu(\omega)$. By Assumption 1.29, it holds that

$$\int u_i(a,\omega) \;\; d\mu(\omega) \; = \; u_{i,1}(a) + u_{i,2}(a) \cdot \bar{\omega}(\mu) + \int u_{i,3}(\omega) \;\; d\mu(\omega) \qquad (1.C.1)$$

for all $\mu \in \Delta\Omega$ and each i.

Now fix any $\mu, \mu' \in \Delta\Omega$ with $\bar{\omega}(\mu) = \bar{\omega}(\mu')$. Condition (1.C.1) implies that $\arg\max_{a\in A} \int u_R(a,\omega) \; d\mu(\omega) = \arg\max_{a\in A} \left(u_{R,1}(a) + u_{R,2}(a) \cdot \bar{\omega}(\mu)\right)$ and thus $\arg\max_{a\in A} \int u_R(a,\omega) \; d\mu(\omega) = \arg\max_{a\in A} \int u_R(a,\omega) \; d\mu'(\omega)$. Take any $a, a' \in A$, and note that by (1.C.1), it follows that $\int u_S(a,\omega) \, d\mu(\omega) \geq \int u_S(a,\omega) \, d\mu(\omega)$ is equivalent to $u_{S,1}(a) + u_{S,2}(a) \cdot \bar{\omega}(\mu) \geq u_{S,1}(a') + u_{S,2}(a') \cdot \bar{\omega}(\mu)$, and thus also equivalent to $\int u_S(a,\omega) \, d\mu'(\omega) \geq \int u_S(a,\omega) \, d\mu'(\omega)$. Finally, for all $\mu \in \mathrm{supp}(\pi)$ and $\omega \in \Omega$, let $v_{i,1}(\mu) \equiv \int u_{i,1}(a) + u_{i,2}(a) \cdot \bar{\omega}(\mu) \, d\sigma_{\mathscr{A}}(a|\pi,\mu)$ and $v_{i,2}(\omega) \equiv u_{i,3}(\omega)$. With that, (1.7) can be easily verified.

Proof of Corollary 1.34.

(i). For any π, let $\bar{\lambda}(\pi)$ be the distribution over posterior means of the state under π. As shown by Kleiner, Moldovanu, and Strack (2021), $\bar{\lambda}(\pi)$ is not an extreme point in the space of distributions over posterior means if π is no bi-pooling policy. Hence, there exist experiments π_1 and π_2 such that $\bar{\lambda}(\pi) = \frac{\bar{\lambda}(\pi_1)+\bar{\lambda}(\pi_2)}{2}$. Note that $\mathrm{supp}(\bar{\lambda}(\pi_1)), \mathrm{supp}(\bar{\lambda}(\pi_2)) \subseteq \mathrm{supp}(\bar{\lambda}(\pi))$. So by Lemma 1.31, π_1 and π_2 are implementable if π is implementable. The rest of the proof follows exactly the same steps as the proof of Theorem 1.25.

(ii). By the construction as in Arieli et al. (2023), either $\bar{\lambda}(\pi_1)$ or $\bar{\lambda}(\pi_2)$ is a Blackwell garbling of $\bar{\lambda}(\pi)$. The rest of the proof follows the same steps as the proof of Theorem 1.19.

Appendix 1.D Proofs: Cheap-Talk With a Perfectly Informed Sender

Proof of Corollary 1.36. The proof works analogously to the proof of Corollary 1.34: For any experiment π that is no bi-partition, there exist π_1, $\pi_2 \in \Pi$ such that $\bar{\lambda}(\pi) = \frac{\bar{\lambda}(\pi_1)+\bar{\lambda}(\pi_2)}{2}$ and for all $i \in \{1,2\}$, and if $\omega \in \mathrm{supp}(\mu)$ for some $\mu \in \pi_i$, there exists some $\mu' \in \pi$ with $\bar{\omega}(\mu) = \bar{\omega}(\mu')$ so that $\omega \in \mathrm{supp}(\mu')$. The latter property ensures that π_1 and π_2 are implementable if π is.

Here is the construction of π_1 and π_2: Fix some Ω' with $\mu_0(\Omega') > 0$ and $|\Omega'| \geq 2$ such that $Y \equiv \bigcap_{\omega\in\Omega'} \{\bar{\omega}(\mu) : \omega \in \mathrm{supp}(\mu) \text{ and } \mathrm{Prob}(\mu|\omega) > 0\} \geq 3$. Choose $\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3 \in Y$. For each $i \in \{1,2,3\}$, let $p_i^* > 0$ (q_i) be the mass of states in (not in) Ω' associated with $\bar{\omega}_i$, and let x_i^* (y_i) be the conditional mean of this mass. Define the map $f : \mathbb{R} \times \mathbb{R}^5 \to \mathbb{R}^5$ as

$$f(p_1, (p_2, p_3, x_1, x_2, x_3)) = \begin{pmatrix} p_1 x_1 + q_1 y_1 - (p_1 + q_1)\bar{\omega}_1 \\ p_2 x_2 + q_2 y_2 - (p_2 + q_2)\bar{\omega}_2 \\ p_3 x_3 + q_3 y_3 - (p_3 + q_3)\bar{\omega}_3 \\ p_1 + p_2 + p_3 - \mu_0(\Omega') \\ p_1 x_1 + p_2 x_2 + p_3 x_3 - \mu_0(\Omega') \int \omega d\mu_0(\omega) \end{pmatrix}.$$

Note that $f(p_1^*, (p_2^*, p_3^*, x_1^*, x_2^*, x_3^*)) = 0$, and f is continuously differentiable. The derivative of $(p_2, p_3, x_1, x_2, x_3) \mapsto f(p_1^*, (p_2, p_3, x_1, x_2, x_3))$ is invertible at $(p_2^*, p_3^*, x_1^*, x_2^*, x_3^*)$: It can be shown that the determinant of the Jacobian matrix is $p_1^* p_2^* p_3^* (\bar{\omega}_3 - \bar{\omega}_2) \neq 0$. By the implicit function theorem, there exist a function $g : \mathbb{R} \to \mathbb{R}^5$ such that $f(p_1, g(p_1)) = 0$ for all p_1 sufficiently close to p_1^*.

Then, one can construct the experiment π_1 (π_2) based on the parameters $(p_1^* + \epsilon, g(p_1^* + \epsilon))$ $((p_1^* - \epsilon, g(p_1^* - \epsilon)))$ for some $\epsilon > 0$ sufficiently close to zero, keeping everything else as in π. Note that π is a Blackwell garbling of π_1.

Appendix 1.E Proofs: Uniform-Quadratic Case

Proof of Lemma 1.37. Solving $\max_{a \in A} \int -(a-\omega)^2 \, d\mu(\omega)$ yields the first-order condition $a = \bar{\omega}(\mu)$. It is the global maximizer by strict concavity of the objective function.

Take any $\mu, \mu' \in \operatorname{supp}(\pi)$, and notice that the incentive compatibility constraints (1.2) of the sender to fully reveal μ instead of misreporting μ' reduce to

$$\begin{aligned} 0 &\le \int -\big(\bar{\omega}(\mu) - \omega - b\big)^2 + \big(\bar{\omega}(\mu') - \omega - b\big)^2 \, d\mu(\omega) \\ &= \int -\big(\bar{\omega}(\mu) - \omega - b\big)^2 + \big(\bar{\omega}(\mu') - \bar{\omega}(\mu) + \bar{\omega}(\mu) - \omega - b\big)^2 d\mu(\omega) \\ &= \int 2\big(\bar{\omega}(\mu') - \bar{\omega}(\mu)\big) \cdot \big(\bar{\omega}(\mu) - \omega - b\big) + \big(\bar{\omega}(\mu') - \bar{\omega}(\mu)\big)^2 d\mu(\omega) \\ &= \big(\bar{\omega}(\mu') - \bar{\omega}(\mu) - 2b\big) \cdot \big(\bar{\omega}(\mu') - \bar{\omega}(\mu)\big). \end{aligned}$$

Hence, either $\bar{\omega}(\mu') - \bar{\omega}(\mu) \ge 2b$ or $\bar{\omega}(\mu') - \bar{\omega}(\mu) \le 0$ must hold. Similarly, the sender has an incentive to truthfully reveal μ' instead of deviating to μ if

$$0 \le \big(\bar{\omega}(\mu) - \bar{\omega}(\mu') - 2b\big)\big(\bar{\omega}(\mu) - \bar{\omega}(\mu')\big),$$

implying that $\bar{\omega}(\mu) - \bar{\omega}(\mu') \ge 2b$ or $\bar{\omega}(\mu) - \bar{\omega}(\mu') \le 0$ must hold. These two incentive compatibility constraints yield $\big|\bar{\omega}(\mu) - \bar{\omega}(\mu')\big| \ge 2b$.

Ex-ante expected payoffs. Notice that

$$\tilde{\mathscr{U}}_R\left(\pi, \tilde{\sigma}_{\mathscr{A}}\right) = \iint -\left(\bar{\omega}(\mu) - \omega\right)^2 \, d\mu\left(\omega\right) \, d\pi\left(\mu\right) = -\sum_{i=1}^{n} p_i \mathrm{Var}\left(\omega | \mu_i\right),$$

where $\mathrm{Var}\left(\omega|\mu_i\right)$ is the conditional distribution of the state given μ_i. These conditional variances take simple analytic forms for bi-pooling policies: If μ_i forms a 1-partition, the state is uniformly distributed on the interval $\left[\bar{\omega}_i - \frac{p_i}{2}, \bar{\omega}_i + \frac{p_i}{2}\right]$, so $\mathrm{Var}\left(\omega|\mu_i\right) = \frac{p_i^2}{12}$. Let $\Omega_i \subseteq \Omega$ be set of states associated with the 1-partition μ_i. To determine the conditional variance given μ_i or μ_{i+1} if μ_i and μ_{i+1} form a 2-partition, let $\Omega_{i,i+1} \subseteq \Omega$ be set of states associated with outcomes μ_i and μ_{i+1} (i.e., if the state lies in $\Omega_{i,i+1}$, the outcome is μ_i or μ_{i+1}). Let $\underline{d}_i = \bar{\omega}_i - \inf\{\Omega_{i,i+1}\}$ and $\bar{d}_{i+1} = \sup\{\Omega_{i,i+1}\} - \bar{\omega}_{i+1}$. Simple algebraic transformations yield

$$\mathrm{Var}\left(\omega|\mu_i \text{ or } \mu_{i+1}\right) = \frac{(p_i + p_{i+1})^2}{12} - \left(\frac{p_i + p_{i+1}}{2} - \underline{d}_i\right)\left(\frac{p_i + p_{i+1}}{2} - \bar{d}_{i+1}\right).$$

Proof of Lemma 1.38. Fix $n \geq 3$ and $b \in \left(\frac{1}{2n}, \frac{1}{2(n-1)}\right]$. First, I show that ignoring all incentive compatibility constraints, uniform partition are the best bi-pooling policies:

Fact 1. *The optimal bi-pooling policy of size $k \in \mathbb{N}$ is the uniform partition of size k.*

Proof. Consider a non-monotone bi-pooling policy with outcomes $\mu_1, \ldots, \mu_k$ and probabilities $p_1, \ldots, p_k$. By non-monotonicity, there is some i so that μ_i and μ_{i+1} form a 2-partition. Consider a new experiment with probabilities $\hat{p}_1, \ldots, \hat{p}_k$ that is constructed from the original experiment by splitting the 2-partition between μ_i and μ_{i+1} into two separate 1-partitions while keeping their weights equal: $\hat{p}_i = p_i$ and $\hat{p}_{i+1} = p_{i+1}$. Note that

$$\begin{aligned}
&\sum_{i=1}^{k} \hat{p}_i \widehat{\mathrm{Var}}\left(\omega|\mu_i\right) - \sum_{i=1}^{k} p_i \mathrm{Var}\left(\omega|\mu_i\right) \\
&\quad = -\left(p_i + p_{i+1}\right)\left(\frac{p_i + p_{i+1}}{2} - \frac{p_i}{2}\right)\left(\frac{p_i + p_{i+1}}{2} - \frac{p_{i+1}}{2}\right) \\
&\qquad + \left(p_i + p_{i+1}\right)\left(\frac{p_i + p_{i+1}}{2} - \underline{d}_i\right)\left(\frac{p_i + p_{i+1}}{2} - \bar{d}_{i+1}\right) \\
&\quad < 0
\end{aligned}$$

since $\underline{d}_i > \frac{p_i}{2}$ and $\bar{d}_{i+1} > \frac{p_{i+1}}{2}$ as μ_i and μ_{i+1} form a 2-partitions. So the optimal bi-pooling policy must be monotone. Now suppose it was not uniform. But then, it holds that

$$\sum_{i=1}^{k} p_i \mathrm{Var}\left(\omega|\mu_i\right) = \sum_{i=1}^{k} \frac{p_i^3}{12} > \sum_{i=1}^{k} \frac{\left(\frac{1}{k}\right)^3}{12},$$

where the last term is the weighted conditional variance of the uniform partition of size k. The inequality follows from convexity of x^3 and $\sum_{i=1}^k p_i = \sum_{i=1}^k \frac{1}{k} = 1$. □

Next, notice that $\sum_{i=1}^k \frac{\left(\frac{1}{k}\right)^3}{12} > \sum_{i=1}^{n-1} \frac{\left(\frac{1}{n-1}\right)^3}{12}$ for all $k < n-1$, that is, the uniform partition of size $n-1$ yields strictly dominates the uniform partition of size $k < n-1$. Moreover, the uniform partition of size $n-1$ is implementable as the IC-constraints are satisfied: The distance between any two adjacent induced actions is $\frac{1}{n-1} \geq 2b$ as $b \leq \frac{1}{2(n-1)}$. Consequently, the optimal experiment is either of size n or the uniform partition of size $n-1$.

It remains to show that under an optimal bi-pooling policy of size n, the outcomes μ_1 and μ_n form 1-partitions and the incentive compatibility constraints of all adjacent induced actions are binding. The proof is completed by the following two facts:

Fact 2. *For any optimal bi-pooling policy, both μ_1 and μ_n are 1-partitions.*

Proof. Suppose μ_1 is no 1-partition, that is, μ_1 and μ_2 form a 2-partition. Hence, it holds that $p_1 + p_2 > 2 \cdot (\bar{\omega}_2 - \bar{\omega}_1)$ and $\max\left\{\underline{d}_1, \bar{d}_2\right\} < \frac{p_1+p_2}{2}$. One obtains that

$$\operatorname{Var}\left(\omega|\mu_1 \text{ or } \mu_2\right) = \frac{\left(p_1 + p_2\right)^2}{12} - \left(\frac{p_1 + p_2}{2} - \underline{d}_1\right) \cdot \left(\frac{p_1 + p_2}{2} - \bar{d}_2\right).$$

Construct a new experiment by decreasing $\bar{\omega}_1$ by some $\epsilon > 0$, decreasing p_1 by $\delta = \frac{p_1 \epsilon}{\bar{\omega}_2 - \bar{\omega}_1 + \epsilon} \in (0, p_1)$ and increasing p_2 by δ , keeping everything else unmodified.[40] Note that all IC-constraints remain satisfied because all posterior means except $\bar{\omega}_1$ are the same as under the original experiment, and $\bar{\omega}_1$ decreases. As long as $\epsilon < \frac{p_1+p_2}{2} - \left(\bar{\omega}_2 - \bar{\omega}_1\right)$, μ_1 and μ_2 continue forming a 2-partition. One gets

$$\widehat{\operatorname{Var}}\left(\omega|\mu_1 \text{ or } \mu_2\right) = \frac{\left(p_1 + p_2\right)^2}{12} - \left(\frac{p_1 + p_2}{2} - \left(\underline{d}_1 - \epsilon\right)\right) \cdot \left(\frac{p_1 + p_2}{2} - \bar{d}_2\right).$$

For all $\epsilon \in \left(0, \frac{p_1+p_2}{2} - \left(\bar{\omega}_2 - \bar{\omega}_1\right)\right)$, this variance is strictly smaller than under the original experiment because

$$\widehat{\operatorname{Var}}\left(\omega|\mu_1 \text{ or } \mu_2\right) - \operatorname{Var}\left(\omega|\mu_1 \text{ or } \mu_2\right) = -\epsilon \cdot \left(\frac{p_1 + p_2}{2} - \bar{d}_2\right) < 0.$$

Since the conditional variance of the state given $\omega \notin \Omega_{1,2}$ and the set $\Omega_{1,2}$ (and thus the probability that $\tilde{\omega} \in \Omega_{1,2}$)) are the same under both experiments, the original bi-pooling policy is not optimal.

The proof for the fact that μ_n must be a 1-partition, too, works analogously. □

40. δ is chosen such that the conditional mean of the state given μ_1 or μ_2 remains unchanged, that is, it satisfies $p_1\bar{\omega}_1 + p_2\bar{\omega}_2 = (p_1 - \delta)(\bar{\omega}_1 - \epsilon) + (p_2 + \delta)\bar{\omega}_2$.

Fact 3. *Under any optimal bi-pooling policy of size $n \in \left(\frac{1}{2b}, \frac{1}{2b}+1\right] \cap \mathbb{N}$, all incentive compatibility constraints are binding.*

Proof. Suppose not, i.e., $\bar{\omega} - \bar{\omega}_{i-1} > 2b$ for some $i \in \{2, \ldots, n\}$. There are four different cases to be considered:

Case 1: (a) μ_{i-1} belongs to a 1-partition, and μ_i belongs to a 2-partition.
(b) μ_{i-1} belongs to a 2-partition, and μ_i belongs to a 1-partition.
Case 2: μ_{i-1} and μ_i belong to different 2-partitions.
Case 3: μ_{i-1} and μ_i belong to the same 2-partition.
Case 4: μ_{i-1} and μ_i both belong to a 1-partition.

Proof of Case 1 (a). : Since μ_i and μ_{i+1} form a 2-partition, $p_i + p_{i+1} > 2\left(\bar{\omega}_{i+1} - \omega_i\right)$ and $\max\left\{\underline{d}_i, \bar{d}_{i+1}\right\} < \frac{p_i+p_{i+1}}{2}$.

Now construct a new experiment by decreasing $\bar{\omega}_i$ by some $\epsilon > 0$, decreasing p_i by $\delta = \frac{p_i \epsilon}{\bar{\omega}_{i+1} - \bar{\omega}_i + \epsilon} \in (0, p_i)$, and increasing p_{i+1} by δ so that all incentive constraints remain satisfied, that is,

$$\left(\bar{\omega}_i - \epsilon\right) - \bar{\omega}_{i-1} \ \geq\ 2b \quad \Leftrightarrow \quad \epsilon \ \leq\ \left(\bar{\omega}_i - \bar{\omega}_{i-1}\right) - 2b,$$

and such that μ_i and μ_{i+1} continue forming a 2-partition, i.e.,

$$p_i + p_{i+1} \ >\ 2 \cdot \left(\bar{\omega}_{i+1} - \left(\bar{\omega}_i - \epsilon\right)\right) \quad \Leftrightarrow \quad \epsilon \ <\ \frac{p_i + p_{i+1}}{2} - \left(\bar{\omega}_{i+1} - \bar{\omega}_i\right).$$

The total decrease of the conditional variance of the state given μ_i or μ_{i+1} is

$$\epsilon \cdot \left(\frac{p_i + p_{i+1}}{2} - \bar{d}_{i+1}\right) \ >\ 0.$$

Proof of Case 1 (b). : The proof of this case works analogously to the proof of Case 1 (a): Increasing $\bar{\omega}_{i-1}$ by $\epsilon \in \left(0, \min\left\{\left(\bar{\omega}_{i-1} - \bar{\omega}_{i-2}\right) - 2b, \frac{p_{i-2}+p_{i-1}}{2} - \left(\bar{\omega}_{i-1} - \bar{\omega}_{i-2}\right)\right\}\right)$, decreasing p_{i-1} by $\delta = \frac{p_{i-1}\epsilon}{\bar{\omega}_{i-1} - \bar{\omega}_{i-2} + \epsilon}$ and increasing p_{i-2} by δ yields a better experiment.

Proof of Case 2. : The proof of this case proceeds exactly as in Case 1 (a).

Proof of Case 3. : Note that $\bar{\omega}_{i-1} - \bar{\omega}_{i-2} = \bar{\omega}_{i+1} - \bar{\omega}_i = 2b$ by Case 1/2. Since μ_{i-1} and μ_i form a 2-partition, one can conclude that $\underline{d}_{i-1} + \bar{d}_i > \bar{\omega}_i - \bar{\omega}_{i-1} > 2b$, implying $\underline{d}_{i-1} > b$ or $\bar{d}_i > b$. As $p_{i-1} + p_i = \underline{d}_{i-1} + \left(\bar{\omega}_i - \bar{\omega}_{i-1}\right) + \bar{d}_i$, it follows that $p_{i-1} + p_i > 4b$ and $\underline{d}_{i-1} + \bar{d}_i > \frac{p_{i-1}+p_i}{2}$.

(i) Let's consider the case where $\underline{d}_{i-1} \leq \bar{d}_i$, so in particular it holds that $\bar{d}_i > b$:

If μ_{i+1} is a 1-partition, one can construct a new experiment by shifting the interval of states $\left(\bar{\omega}_i + \bar{d}_i - \epsilon, \bar{\omega}_i + \bar{d}_i\right)$ from outcomes μ_{i-1} or μ_i to outcome μ_{i+1} for some $\epsilon > 0$. Since μ_{i+1} is a 1-partition, $\bar{\omega}_{i+1}$ decreases by exactly $\frac{\epsilon}{2}$. Hence, $\bar{\omega}_i$ must decrease by at least $\frac{\epsilon}{2}$ in order to fulfill the ith incentive compatibility constraint. For concreteness, suppose that $\bar{\omega}_i$ decreases by exactly $\frac{\epsilon}{2}$ as well. To satisfy the $(i-1)$th incentive compatibility constraint, it is necessary that

$$\left(\bar{\omega}_i - \frac{\epsilon}{2}\right) - \bar{\omega}_{i-1} \geq 2b \quad \Leftrightarrow \quad \epsilon \leq \frac{\bar{\omega}_i - \bar{\omega}_{i-1}}{2} - b.$$

One obtains that

$$\begin{aligned}\sum_{j=i-1}^{i+1} \hat{p}_j \widehat{\mathrm{Var}}\left(\omega|\mu_j\right) = & \frac{\left(p_{i-1}+p_i-\epsilon\right)^3}{12} + \frac{\left(p_{i+1}+\epsilon\right)^3}{12} - \left(p_{i-1}+p_i-\epsilon\right) \\ & \cdot \left(\frac{p_{i-1}+p_i-\epsilon}{2} - \underline{d}_{i-1}\right) \cdot \left(\frac{p_{i-1}+p_i-\epsilon}{2} - \left(\bar{d}_i - \frac{\epsilon}{2}\right)\right).\end{aligned}$$

At $\epsilon = 0$, the derivative with respect to ϵ is

$$-\frac{\left(p_{i-1}+p_i\right)^2}{4} + \frac{p_{i+1}^2}{4} + \left(\frac{p_{i-1}+p_i}{2} - \bar{d}_i\right) \cdot \left(p_{i-1}+p_i - \underline{d}_{i-1}\right). \tag{1.E.1}$$

Notice that $\underline{d}_{i-1} > \frac{p_{i-1}+p_i}{2} - \bar{d}_i$ holds due to the fact that μ_{i-1} and μ_i are a 2-partition, and that $p_{i+1} = 2 \cdot \left(2b - \bar{d}_i\right)$ because μ_{i+1} is a 1-partition, which implies $p_{i+1} < 2b$ as $\bar{d}_i > b$. Consequently, the derivative term at $\epsilon = 0$ (1.E.1) is strictly negative as

$$\begin{aligned} & -\frac{\left(p_{i-1}+p_i\right)^2}{4} + \frac{p_{i+1}^2}{4} + \left(\frac{p_{i-1}+p_i}{2} - \bar{d}_i\right) \cdot \left(p_{i-1}+p_i - \underline{d}_{i-1}\right) \\ & < -\frac{\left(p_{i-1}+p_i\right)^2}{4} + \frac{p_{i+1}^2}{4} + \left(\frac{p_{i-1}+p_i}{2} - \bar{d}_i\right) \cdot \left(\frac{p_{i-1}+p_i}{2} + \bar{d}_i\right) \\ & = \frac{p_{i+1}^2}{4} - \bar{d}_i^2 < \frac{(2b)^2}{4} - b^2 = 0. \end{aligned}$$

Since the derivative is continuous in ϵ, it is thus strictly negative on some non-empty open set around zero. So for any $\epsilon > 0$ sufficiently close to zero, the conditional variance of the state given μ_{i-1} or μ_i or μ_{i+1} is strictly smaller under the new experiment compared to the original bi-pooling policy (which corresponds to the case when $\epsilon = 0$), which therefore cannot be optimal.

If μ_{i+1} and μ_{i+2} form a 2-partition and the (i+1)th incentive constraint is binding, i.e., $\bar{\omega}_{i+2} - \bar{\omega}_{i+1} = 2b$, construct a new experiment by shifting the interval $(\bar{\omega}_i + \bar{d}_i - \epsilon, \bar{\omega}_i + \bar{d}_i)$ from outcomes μ_{i-1} or μ_i to outcomes μ_{i+1} or μ_{i+2} for some $\epsilon > 0$ leaving $\bar{\omega}_i$ and $\bar{\omega}_{i+1}$ unchanged. As long as

$$\begin{aligned} & \underline{d}_{i-1} + \bar{d}_i - \epsilon \geq \frac{p_{i-1}+p_i-\epsilon}{2} \\ \Leftrightarrow \quad & \epsilon \leq 2 \cdot \left(\underline{d}_{i-1} + \bar{d}_i - \frac{p_{i-1}+p_i}{2}\right) = 2 \cdot \left(\underline{d}_{i-1} + \bar{d}_i\right) - \left(p_{i-1}+p_i\right), \end{aligned}$$

μ_{i+1} and μ_{i+2} continue being a 2-partition. One obtains that

$$\begin{aligned}\sum_{j=i-1}^{i+2} \hat{p}_j \widehat{\mathrm{Var}}\left(\omega|\mu_j\right) = & \frac{\left(p_{i-1}+p_i-\epsilon\right)^3}{12} - \left(p_{i-1}+p_i-\epsilon\right) \\ & \cdot \left(\frac{p_{i-1}+p_i-\epsilon}{2} - \underline{d}_{i-1}\right) \cdot \left(\frac{p_{i-1}+p_i-\epsilon}{2} - \left(\bar{d}_i - \epsilon\right)\right) \\ & + \frac{\left(p_{i+1}+p_{i+2}+\epsilon\right)^3}{12} - \left(p_{i+1}+p_{i+2}+\epsilon\right) \\ & \cdot \left(\frac{p_{i+1}+p_{i+2}+\epsilon}{2} - \left(\underline{d}_{i+1}+\epsilon\right)\right) \cdot \left(\frac{p_{i+1}+p_{i+2}+\epsilon}{2} - \bar{d}_{i+2}\right).\end{aligned}$$

At $\epsilon = 0$, the derivative with respect to ϵ becomes

$$\begin{aligned}& -\bar{d}_i \cdot \left(\left(p_{i-1}+p_i\right) - \underline{d}_{i-1}\right) + \underline{d}_{i+1} \cdot \left(\left(p_{i+1}+p_{i+2}\right) - \bar{d}_{i+2}\right) \\ & \quad = -\bar{d}_i \cdot \left(\left(a_i - a_{i-1}\right) + \bar{d}_i\right) + \underline{d}_{i+1} \cdot \left(\underline{d}_{i+1} + \left(a_{i+2} - a_{i+1}\right)\right) \\ & \quad < -b \cdot (2b+b) + b \cdot (b+2b) = 0.\end{aligned}$$

By continuity of the derivative term in ϵ, there exists some non-empty open interval around zero such that the derivative is strictly negative on this interval. So for any $\epsilon > 0$ sufficiently close to zero, the conditional variance of the state given μ_{i-1}, μ_i, μ_{i+1} or μ_{i+2} is strictly smaller than under the original experiment.

As a consequence, the original experiment can only be optimal if μ_{i+1} and μ_{i+2} form a 2-partition and $\bar{\omega}_{i+2} - \bar{\omega}_{i+1} > 2b$. Since $\underline{d}_{i+1} = \left(\bar{\omega}_{i+1} - \bar{\omega}_i\right) - \bar{d}_i < 2b - b = b$, this implies $\bar{d}_{i+2} > b$. So by applying the same reasoning as above, this can only be optimal if μ_{i+3} and μ_{i+4} form a 2-partition and $\bar{\omega}_{i+4} - \bar{\omega}_{i+3} > 2b$. Iterating forward, it is thus necessary that for all $j \in \{i, i+2, \ldots, n-2, n\}$, μ_{j-1} and μ_j form a 2-partition and $\bar{\omega}_j - \bar{\omega}_{j-1} > 2b$. So in particular, $n - i$ must be even. However, since μ_n forms a 1-partition by Fact 2, this is not possible under an optimal bi-pooling policy.

(ii) The proof for the case where $\underline{d}_{i-1} + > \bar{d}_i$ works analogously: One can show that μ_{i-2} cannot form a 1-partition, but for all $j \in \{2, 4, \ldots, i-2, i\}$, μ_j must belong to a 2-partition together with μ_{j-1} such that the $(j-1)$th incentive compatibility constraint is binding—implying that i must be even. However, this contradicts the fact that μ_1 forms a 1-partition by Fact 2.

Proof of Case 4. : If both μ_i and μ_{i-1} are 1-partitions, it follows from $\bar{\omega}_i - \bar{\omega}_{i-1} > 2b$ that $p_{i-1} + p_i = 2 \cdot \left(\bar{\omega}_i - \bar{\omega}_{i-1}\right) > 4b$.

Let's focus on the case $p_i \geq p_{i-1}$. Consequently, it holds that $p_i > 2b$.

Suppose first that $i = n$, i.e., $p_n > 2b$. Since μ_1 and μ_n form a 1-partition, respectively, it holds that $\underline{d}_1 = \frac{p_1}{2}$ and $\bar{d}_n = \frac{p_n}{2}$. Hence, one obtains that

$$\frac{p_1 + p_n}{2} = \underline{d}_1 + \bar{d}_n = 1 - \sum_{j=2}^{n} \bar{\omega}_j - \bar{\omega}_{j-1} \leq 1 - (n-1) \cdot 2b < 2b,$$

where the first, weak inequality holds true due to the fact that all incentive compatibility constraints are satisfied, and the second, strong inequality follows from the fact that $b > \frac{1}{2n}$. One thus gets that $p_1 < 2b < p_n$.

Consider now the following alternative experiment: Increase p_1 by some $\epsilon > 0$ and decrease p_n by the same ϵ such that

$$\frac{p_{n-1} + p_n - \epsilon}{2} \;\geq\; 2b \quad\Leftrightarrow\quad \epsilon \;\leq\; 4b - \left(p_{n-1} + p_n\right).$$

One obtains that

$$\sum_{j\in\{1,n\}} \hat{p}_j \widehat{\mathrm{Var}}\left(\omega|\mu_j\right) \;=\; \frac{\left(p_1 + \epsilon\right)^3}{12} + \frac{\left(p_n - \epsilon\right)^3}{12}.$$

Its derivative with respect to ϵ is $\frac{p_1^2 - p_n^2}{4} < 0$ at $\epsilon = 0$. So the derivative is strictly negative on some non-empty, open interval around $\epsilon = 0$. Hence, the conditional variance of the state given μ_1 or μ_n is strictly smaller under the alternative experiment for any $\epsilon > 0$ sufficiently close to zero, implying that the original bi-pooling policy cannot be optimal. Consequently, it must be that $i < n$.

Suppose μ_{i+1} is a 1-partition. If $p_{i+1} \geq p_i$, one gets $\bar{\omega}_{i+1} - \bar{\omega}_i = \frac{p_{i+1}+p_i}{2} > \frac{2b+2b}{2} = 2b$. But then, one can find a better experiment by decreasing p_i by some $\epsilon > 0$ and increasing $\min\{p_1, p_n\} < 2b$ by the same ϵ so that all incentive compatibility constraints remain satisfied, i.e., $p_{i+1} + p_i - \epsilon \geq 4b$ and $p_i + p_{i-1} - \epsilon \geq 4b$, or, equivalently, $\epsilon \leq 4b - p_i - p_{i+1}$. Note that

$$\sum_{j\in\{1,i,n\}} \hat{p}_j \widehat{\mathrm{Var}}\left(\omega|\mu_j\right) \;=\; \frac{\left(\min\{p_1, p_n\} + \epsilon\right)^3}{12} + \frac{\left(p_i - \epsilon\right)^3}{12}.$$

The derivative with respect to ϵ is $\frac{\min\{p_1,p_n\}^2 - p_i^2}{4} < 0$ at $\epsilon = 0$, implying that the original experiment is not optimal.

Hence, it must be that $p_{i+1} < p_i$. But then, there is a better experiment under which p_{i+1} is increased by some $\epsilon > 0$ while p_i is reduced by the same ϵ such that all incentive compatibility constraints remain valid, i.e., $p_{i-1} + p_i - \epsilon \geq 4b$, and such that the conditional variance of the state given μ_i or μ_{i+1} shrinks because the derivative of

$$\sum_{j\in\{i,i+1\}} \hat{p}_j \widehat{\mathrm{Var}}\left(\omega|\mu_j\right) \;=\; \frac{\left(p_i - \epsilon\right)^3}{12} + \frac{\left(p_{i+1} + \epsilon\right)^3}{12}$$

with respect to ϵ is strictly negative at $\epsilon = 0$.

Hence, μ_{i+1} must belong to a 2-partition. Let j be the smallest integer larger than $i + 1$ such that μ_j is a 1-partition. This is well-defined as μ_n is a 1-partition by Fact 2. Then, $j - (i + 1)$ is even, and for all $k \in \{1, \ldots, \frac{j-(i+1)}{2}\}$, μ_{2k+i-1} and μ_{2k+i}

form a 2-partition. Moreover, one obtains that $\bar{\omega}_l - \bar{\omega}_{l-1} = 2b$ for all $l \in \{i+1, \dots, j\}$ by Cases 1-3. Consider the following alternative experiment: For some $\epsilon \in (0, b)$, shift the interval $\left(\sup\{\Omega_i\} - \epsilon, \sup\{\Omega_i\}\right)$ from outcome μ_i to outcomes μ_{i+1} or μ_{i+2}, for each $k \in \{1, \dots, \frac{j-(i+1)}{2} - 1\}$, shift $\left(\sup\{\Omega_{2k+i-1,2k+i}\} - \epsilon, \sup\{\Omega_{2k+i-1,2k+i}\}\right)$ from $\Omega_{2k+i-1,2k+i}$ to $\Omega_{2k+i+1,2k+i+2}$, and shift $\left(\sup\{\Omega_{j-2,j-1}\} - \epsilon, \sup\{\Omega_{j-2,j-1}\}\right)$ from $\Omega_{j-2,j-1}$ to Ω_j such that the posterior means $\bar{\omega}_i, \bar{\omega}_{i+1}, \dots, \bar{\omega}_{j-1}, \bar{\omega}_j$ decrease by $\frac{\epsilon}{2}$, respectively.[41] The fact that both $\bar{\omega}_i$ and $\bar{\omega}_j$ decrease by $\frac{\epsilon}{2}$ yields that μ_i and μ_j remain 1-partitions. Furthermore, the construction ensures that μ_{2k+i-1} and μ_{2k+i} continue forming a 2-partition for all $k \in \{1, \dots, \frac{j-(i+1)}{2}\}$. Additionally, as long as $\bar{\omega}_i - \bar{\omega}_{i-1} - \frac{\epsilon}{2} \geq 2b$, all incentive compatibility constraints remain valid. Besides, note that $\underline{d}_{2k+i-1} < b$ and $\bar{d}_{2k+i} > b$ holds for all $k \in \{1, \dots, \frac{j-(i+1)}{2}\}$, which can be shown by induction on k: For $k = 1$, this is true because

$$\underline{d}_{i+1} = \left(\bar{\omega}_{i+1} - \bar{\omega}_i\right) - \bar{d}_i = 2b - \frac{p_i}{2} < 2b - \frac{2b}{2} = b,$$

and thus

$$\bar{d}_{i+2} = p_{i+1} + p_{i+2} - \left(\bar{\omega}_{i+2} - \bar{\omega}_{i+1}\right) - \underline{d}_{i+1} > 4b - 2b - b = b.$$

So for $k > 1$, one obtains that

$$\underline{d}_{2k+i-1} = \left(\bar{\omega}_{2(k-1)+i+1} - \bar{\omega}_{2(k-1)+i}\right) - \bar{d}_{2(k-1)+i} = 2b - \hat{\bar{d}}_{2(k-1)+i} < b,$$

as $\bar{d}_{2(k-1)+i} > b$ by the induction hypothesis $(k-1)$, and hence

$$\bar{d}_{2k+i} = p_{2k+i-1} + p_{2k+i} - \left(\bar{\omega}_{2k+i} - \bar{\omega}_{2k+i-1}\right) - \underline{d}_{2k+i-1} > 4b - 2b - b = b.$$

For any $k \in \{1, \dots, \frac{j-(i+1)}{2}\}$, one obtains

$$\begin{aligned} \sum_{l=2k+i-1}^{2k+i} \hat{p}_l \widehat{\mathrm{Var}}\left(\omega | \mu_l\right) = \; & \frac{\left(p_{2k+i-1} + p_{2k+i}\right)^3}{12} - \left(p_{2k+i-1} + p_{2k+i}\right) \\ & \cdot \left(\frac{p_{2k+i-1} + p_{2k+i}}{2} - \left(\underline{d}_{2k+i-1} + \frac{\epsilon}{2}\right)\right) \\ & \cdot \left(\frac{p_{2k+i-1} + p_{2k+i}}{2} - \left(\bar{d}_{2k+i} - \frac{\epsilon}{2}\right)\right). \end{aligned}$$

At $\epsilon = 0$, the derivative with respect to ϵ is

$$\frac{p_{2k+i-1} + p_{2k+i}}{2} \cdot \left(\underline{d}_{2k+i-1} - \bar{d}_{2k+i}\right) < \frac{p_{2k+i-1} + p_{2k+i}}{2} \cdot (b - b) = 0.$$

41. Choosing $\epsilon < b$ ensures that intervals are shifted from one 1- or 2-partition to the next one only, that is, $\left(\sup\{\Omega_i\} - \epsilon, \sup\{\Omega_i\}\right) \subseteq \Omega_i$ and $\left(\sup\{\Omega_{2k+i-1,2k+i}\} - \epsilon, \sup\{\Omega_{2k+i-1,2k+i}\}\right) \subseteq \Omega_{2k+i-1,2k+i}$ for every $k \in \{1, \dots, \frac{j-(i+1)}{2}\}$.

Moreover, note that

$$\sum_{l\in\{i,j\}} \hat{p}_l \widehat{\mathrm{Var}}\left(\omega|\mu_l\right) = \frac{\left(p_i-\epsilon\right)^3}{12} + \frac{\left(p_j+\epsilon\right)^3}{12},$$

and its derivative with respect to ϵ is $-\frac{p_i^2-p_j^2}{4} < 0$ at $\epsilon = 0$ because $p_i > 2b$ and $p_j = 2\underline{d}_j = 2\left(\bar{\omega}_j - \bar{\omega}_{j-1} - \bar{d}_{j-1}\right) < 2(2b-b) = 2b$.

Therefore, one can conclude by the same reasoning as in the above three cases that there is a better experiment for some $\epsilon > 0$ sufficiently close to zero. □

Proof of Lemma 1.39, Lemma 1.40 and Theorem 1.41.

Optimal Partial Bi-pooling Policies. Consider the following auxiliary problem: Fix $b > 0$, $j \in \mathbb{N}$ and $(\underline{\omega}, \bar{\omega}) \subset [0,1]$. The aim is to find an optimal *partial bi-pooling policy* of size j on the interval $(\underline{\omega}, \bar{\omega})$. I characterize the optimal partial bi-pooling policies with $\bar{\omega}_i - \bar{\omega}_{i-1} = 2b$ for all $i \in \{2,\dots,j\}$, $\underline{d}_1 \equiv \bar{\omega}_1 - \underline{\omega} \in (0,2b)$ and $\bar{d}_j \equiv \bar{\omega} - \bar{\omega}_j \in (0,2b)$.

Optimal Partial Bi-pooling Policy Given $\underline{d}_1$ and $\bar{d}_j$. I characterize the optimal partial bi-pooling policy for any fixed $\underline{d}_1 \in (0,2b)$ and $\bar{d}_j \in (0,2b)$ and for any $j \in \mathbb{N}$: I determine under which conditions Ω_1 forms a 1-partition or a 2-partition with Ω_2. Then, I compute the optimal conditional variance, which turns out to depend on $\underline{d}_1$, $\bar{d}_j$ and j only. Let $v^* : (0,2b) \times (0,2b) \times \mathbb{N} \to \mathbb{R}$ be defined by

$$v^*(\underline{d}_1, \bar{d}_j, j) = \left(\underline{d}_1 + 2b(j-1) + \bar{d}_j\right) \cdot \mathrm{Var}(\omega|\omega \in (\underline{\omega}, \bar{\omega})).$$

***j* = 1:** Ω_1 forms a 1-partition, so $\underline{d}_1 = \bar{d}_1$. Therefore, $v^*(\underline{d}_1, \bar{d}_1, 1) = \frac{(\underline{d}_1+\bar{d}_1)^3}{12}$.

***j* = 2:** Ω_1 and Ω_2 form a 2-partition, so $\underline{d}_1 + \bar{d}_2 \geq 2b$. One obtains

$$v^*(\underline{d}_1, \bar{d}_2, 2) = \frac{\underline{d}_1^3 + \bar{d}_2^3}{3} + b \cdot (\underline{d}_1^2 + \bar{d}_2^2) - \frac{4b^3}{3}.$$

***j* = 3:** Ω_1 forms a 1-partition iff $\underline{d}_1 \leq \bar{d}_3$. Ω_1 and Ω_2 form a 2-partition iff $\underline{d}_1 \geq \bar{d}_3$. One obtains

$$\begin{aligned} v^*(\underline{d}_1, \bar{d}_3, 3) &= \begin{cases} v^*(\underline{d}_1, \underline{d}_1, 1) + v^*(2b - \underline{d}_1, \bar{d}_3, 2) & \text{, if } \underline{d}_1 \leq \bar{d}_3 \\ v^*(\underline{d}_1, 2b - \bar{d}_3, 2) + v^*(\bar{d}_3, \bar{d}_3, 1) & \text{, if } \underline{d}_1 \geq \bar{d}_3 \end{cases} \\ &= \begin{cases} \frac{\underline{d}_1^3 + \bar{d}_3^3}{3} + 3b\underline{d}_1^2 + b\bar{d}_3^2 - 8b^2\underline{d}_1 + \frac{16b^3}{3} & \text{, if } \underline{d}_1 \leq \bar{d}_3 \\ \frac{\underline{d}_1^3 + \bar{d}_3^3}{3} + b\underline{d}_1^2 + 3b\bar{d}_3^2 - 8b^2\bar{d}_3 + \frac{16b^3}{3} & \text{, if } \underline{d}_1 \geq \bar{d}_3 \end{cases} \end{aligned}$$

Lemma 1.50. *If $j \geq 4$, any optimal partial bi-pooling policy on $(\underline{\omega}, \bar{\omega})$ satisfies:*
(i) If j is even, Ω_1 forms a 1-partition iff $\underline{d}_1 \leq \min\{\frac{(j-2)b}{j-3}, \max\{b, 2b - \bar{d}_j\}\}$. Ω_1 and Ω_2 are a 2-partition iff $\underline{d}_1 \geq \min\{\frac{(j-2)b}{j-3}, \max\{b, 2b - \bar{d}_j\}\}$. The optimal value is

$$v^*(\underline{d}_1, \bar{d}_j, j) =$$

$$\begin{cases} \frac{\underline{d}_1^3+\bar{d}_j^3}{3} + 3b\left(\underline{d}_1^2 + \bar{d}_j^2\right) - 8b^2\left(\underline{d}_1 + \bar{d}_j\right) + \frac{2(j+14)b^3}{3} & , \underline{d}_1 \leq b \wedge \bar{d}_j \leq b \\ \frac{\underline{d}_1^3+\bar{d}_j^3}{3} + (2j-5)b\underline{d}_1^2 + 3b\bar{d}_j^2 - 4(j-2)b^2\underline{d}_1 - 8b^2\bar{d}_j + \frac{(8j+4)b^3}{3} & , \underline{d}_1 + \bar{d}_j \leq 2b \wedge \frac{(j-2)b}{j-3} \geq \underline{d}_1 \geq b \\ \frac{\underline{d}_1^3+\bar{d}_j^3}{3} + 3b\underline{d}_1^2 + (2j-5)b\bar{d}_j^2 - 8b^2\underline{d}_1 - 4(j-2)b^2\bar{d}_j + \frac{(8j+4)b^3}{3} & , \underline{d}_1 + \bar{d}_j \leq 2b \wedge \frac{(j-2)b}{j-3} \geq \bar{d}_j \geq b \\ \frac{\underline{d}_1^3+\bar{d}_j^3}{3} + b\underline{d}_1^2 + 3b\bar{d}_j^2 - 8b^2\bar{d}_j + \left(\frac{2(j+5)}{3} - \frac{2}{j-3}\right)b^3 & , \underline{d}_1 \geq \frac{(j-2)b}{j-3} \wedge \bar{d}_j \leq \frac{(j-4)b}{j-3} \\ \frac{\underline{d}_1^3+\bar{d}_j^3}{3} + 3b\underline{d}_1^2 + b\bar{d}_j^2 - 8b^2\underline{d}_1 + \left(\frac{2(j+5)}{3} - \frac{2}{j-3}\right)b^3 & , \underline{d}_1 \leq \frac{(j-4)b}{j-3} \wedge \bar{d}_j \geq \frac{(j-2)b}{j-3} \\ \frac{\underline{d}_1^3+\bar{d}_j^3}{3} + (2j-3)b\underline{d}_1^2 + b\bar{d}_j^2 - 4(j-2)b^2\underline{d}_1 + \frac{4(2j-5)b^3}{3} & , \underline{d}_1 + \bar{d}_j \geq 2b \wedge b \geq \underline{d}_1 \geq \frac{(j-4)b}{j-3} \\ \frac{\underline{d}_1^3+\bar{d}_j^3}{3} + b\underline{d}_1^2 + (2j-3)b\bar{d}_j^2 - 4(j-2)b^2\bar{d}_j + \frac{4(2j-5)b^3}{3} & , \underline{d}_1 + \bar{d}_j \geq 2b \wedge b \geq \bar{d}_j \geq \frac{(j-4)b}{j-3} \\ \frac{\underline{d}_1^3+\bar{d}_j^3}{3} + b\cdot\left(\underline{d}_1^2 + \bar{d}_j^2\right) + \frac{2(j-4)b^3}{3} & , \underline{d}_1 \geq b \wedge \bar{d}_j \geq b \end{cases}$$

(ii) If j is odd, Ω_1 forms a 1-partition iff $\underline{d}_1 \leq \min\{\frac{j-1}{j-2} \cdot b, \max\{b, \bar{d}_j\}\}$. Ω_1 and Ω_2 are a 2-partition iff $\underline{d}_1 \geq \min\{\frac{j-1}{j-2} \cdot b, \max\{b, \bar{d}_j\}\}$. The optimal value is

$$v^*(\underline{d}_1, \bar{d}_j, j) =$$

$$\begin{cases} \frac{\underline{d}_1^3+\bar{d}_j^3}{3} + 3b\left(\underline{d}_1^2 + \bar{d}_j^2\right) - 8b^2\left(\underline{d}_1 + \bar{d}_j\right) + \left(\frac{2(j+14)}{3} - \frac{2}{j-4}\right)b^3 & , \underline{d}_1 \leq \frac{(j-5)b}{j-4} \wedge \bar{d}_j \leq \frac{(j-5)b}{j-4} \\ \frac{\underline{d}_1^3+\bar{d}_j^3}{3} + (2j-5)b\underline{d}_1^2 + 3b\bar{d}_j^2 - 4(j-3)b^2\underline{d}_1 - 8b^2\bar{d}_j + \frac{8(j-1)b^3}{3} & , \underline{d}_1 \geq \bar{d}_j \wedge b \geq \underline{d}_1 \geq \frac{(j-5)b}{j-4} \\ \frac{\underline{d}_1^3+\bar{d}_j^3}{3} + 3b\underline{d}_1^2 + (2j-5)b\bar{d}_j^2 - 8b^2\underline{d}_1 - 4(j-3)b^2\bar{d}_j + \frac{8(j-1)b^3}{3} & , \bar{d}_j \geq \underline{d}_1 \wedge b \geq \bar{d}_j \geq \frac{(j-5)b}{j-4} \\ \frac{\underline{d}_1^3+\bar{d}_j^3}{3} + b\underline{d}_1^2 + 3b\bar{d}_j^2 - 8b^2\bar{d}_j + \frac{2(j+5)b^3}{3} & , \underline{d}_1 \geq b \geq \bar{d}_j \\ \frac{\underline{d}_1^3+\bar{d}_j^3}{3} + 3b\underline{d}_1^2 + b\bar{d}_j^2 - 8b^2\underline{d}_1 + \frac{2(j+5)b^3}{3} & , \bar{d}_j \geq b \geq \underline{d}_1 \\ \frac{\underline{d}_1^3+\bar{d}_j^3}{3} + b\underline{d}_1^2 + (2j-3)b\bar{d}_j^2 - 4(j-1)b^2\bar{d}_j + \frac{8(j-1)b^3}{3} & , \underline{d}_1 \geq \bar{d}_j \wedge \frac{(j-1)b}{j-2} \geq \bar{d}_j \geq b \\ \frac{\underline{d}_1^3+\bar{d}_j^3}{3} + (2j-3)b\underline{d}_1^2 + b\bar{d}_j^2 - 4(j-1)b^2\underline{d}_1 + \frac{8(j-1)b^3}{3} & , \bar{d}_j \geq \underline{d}_1 \wedge \frac{(j-1)b}{j-2} \geq \underline{d}_1 \geq b \\ \frac{\underline{d}_1^3+\bar{d}_j^3}{3} + b\cdot\left(\underline{d}_1^2 + \bar{d}_j^2\right) + \frac{2(j-1)(j-5)b^3}{3(j-2)} & , \underline{d}_1 \geq \frac{(j-1)b}{j-2} \wedge \bar{d}_j \geq \frac{(j-1)b}{j-2} \end{cases}$$

Proof. This result is proven by induction on j with two base cases ($j = 4, 5$) and two induction steps, one for j odd and the other for j even:

Base case 1: $j = 4$. Suppose Ω_1 forms a 1-partition. Using the result of the case $j = 3$, this experiment is uniquely determined: Ω_2 forms a 1-partition if and only if $2b - \underline{d}_1 \leq \bar{d}_4$, and Ω_2 and Ω_3 form a 2-partition if and only if $2b - \underline{d}_1 \geq \bar{d}_4$.

If Ω_1 and Ω_2 form a 2-partition, the optimal value of a partial bi-pooling policy on $\omega \in (\underline{\omega}, \bar{\omega})$ is given by

$$\min_{d} \quad v^*(\underline{d}_1, d, 2) + v^*(2b - d, \bar{d}_4, 2) \quad \text{s.t.} \quad \underline{d}_1 + d \geq 2b \text{ and } (2b - d) + \bar{d}_4 \geq 2b.$$

The constraint set is non-empty if and only if $2b - \underline{d}_1 \leq \bar{d}_4$. The solution of the optimization problem is

$$d^* = \begin{cases} b & , \text{ if } \bar{d}_4 \geq b \geq 2b - \underline{d}_1 \\ \bar{d}_4 & , \text{ if } b \geq \bar{d}_4 \geq 2b - \underline{d}_1 \\ 2b - \underline{d}_1 & , \text{ if } \bar{d}_4 \geq 2b - \underline{d}_1 \geq b \end{cases}$$

One obtains

$$\begin{aligned} & v^*(\underline{d}_1, \bar{d}_4, 4) \\ & = \begin{cases} v^*(\underline{d}_1, \underline{d}_1, 1) + v^*(2b - \underline{d}_1, \bar{d}_4, 3) & , \text{if } 2b - \underline{d}_1 \geq \bar{d}_4 \\ v^*(\underline{d}_1, d^*, 2) + v^*(2b - d^*, \bar{d}_4, 2) & , \text{if } 2b - \underline{d}_1 \leq \bar{d}_4 \end{cases} \\ & = \begin{cases} \frac{\underline{d}_1^3 + \bar{d}_4^3}{3} + 3b\left(\underline{d}_1^2 + \bar{d}_4^2\right) - 8b^2\left(\underline{d}_1 + \bar{d}_4\right) + 12b^3 & , \text{if } \underline{d}_1 + \bar{d}_4 \leq 2b \\ \frac{\underline{d}_1^3 + \bar{d}_4^3}{3} + 5b\underline{d}_1^2 + b\bar{d}_4^2 - 8b^2\underline{d}_1 + 4b^3 & , \text{if } b \geq \underline{d}_1 \geq 2b - \bar{d}_4 \\ \frac{\underline{d}_1^3 + \bar{d}_4^3}{3} + b\underline{d}_1^2 + 5b\bar{d}_4^2 - 8b^2\bar{d}_4 + 4b^3 & , \text{if } b \geq \bar{d}_4 \geq 2b - \underline{d}_1 \\ \frac{\underline{d}_1^3 + \bar{d}_4^3}{3} + b\left(\underline{d}_1^2 + \bar{d}_4^2\right) & , \text{if } \underline{d}_1 \geq b \text{ and } \bar{d}_4 \geq b \end{cases} \end{aligned}$$

Ω_1 forms a 1-partition in the first two cases: $\underline{d}_1 + \bar{d}_4 \leq 2b$ or $b \geq \underline{d}_1 \geq 2b - \bar{d}_4$.

Base case 2: $j = 5$. Suppose Ω_1 forms a 1-partition. Using the result of case $j = 4$, this experiment is uniquely determined: Ω_2 forms a 1-partition if and only if $2b - \underline{d}_1 \leq \max\{b, 2b - \bar{d}_j\}$, or, equivalently, $\underline{d}_1 \geq \min\{b, \bar{d}_5\}$. Ω_2 and Ω_3 form a 2-partition if and only if $\underline{d}_1 \leq \min\{b, \bar{d}_5\}$.

If Ω_1 and Ω_2 form a 2-partition, the optimal value of a partial bi-pooling policy on $\omega \in (\underline{\omega}, \bar{\omega})$ is given by

$$\min_d \quad v^*(\underline{d}_1, d, 2) + v^*(2b - d, \bar{d}_5, 3) \quad \text{s.t.} \quad \underline{d}_1 + d \geq 2b.$$

The solution to the optimization problem is

$$d^* = \begin{cases} b & , \text{ if } \underline{d}_1 \geq b \geq \bar{d}_5 \\ \frac{2b}{3} & , \text{ if } \underline{d}_1 \geq \frac{4b}{3} \text{ and } \bar{d}_5 \geq \frac{4b}{3} \\ 2b - \bar{d}_5 & , \text{ if } \underline{d}_1 \geq \bar{d}_5 \text{ and } \frac{4b}{3} \geq \bar{d}_5 \geq b \\ 2b - \underline{d}_1 & , \underline{d}_1 \leq \min\{\frac{4b}{3}, \max\{b, \bar{d}_5\}\} \end{cases}$$

If $\underline{d}_1 \leq \min\{\frac{4b}{3}, \max\{b, \bar{d}_5\}\}$, the best experiment under which Ω_1 and Ω_2 form a 2-partition is essentially one under which Ω_1 and Ω_2 form two separate 1-partitions as the solution to the optimization problem is $d^* = 2b - \underline{d}_1$. Consequently, the globally optimal experiment must be the one under which Ω_1 is a 1-partition.

Notice that $\min\{b,\bar{d}_5\} \le \min\{\frac{4b}{3},\max\{b,\bar{d}_5\}\}$. If $\underline{d}_1 \ge \min\{\frac{4b}{3},\max\{b,\bar{d}_5\}\}$, implying that $\underline{d}_1 \ge \min\{b,\bar{d}_5\}$, the experiment under which Ω_1 is a 1-partition belongs to the constraint set of the minimization problem. But it is not optimal as $d^* \neq 2b-\underline{d}_1$. The globally optimal experiment is thus one under which Ω_1 belongs to a 2-partition. One obtains

$$
\begin{aligned}
&v^*(\underline{d}_1,\bar{d}_5,5)\\
&= \begin{cases} v^*(\underline{d}_1,\underline{d}_1,1)+v^*(2b-\underline{d}_1,\bar{d}_5,4) & ,\text{if } \underline{d}_1 \le \min\{\frac{4b}{3},\max\{b,\bar{d}_j\}\} \\ v^*(\underline{d}_1,d,2)+v^*(2b-d,\bar{d}_5,3) & ,\text{if } \underline{d}_1 \ge \min\{\frac{4b}{3},\max\{b,\bar{d}_j\}\} \end{cases}\\
&= \begin{cases} \frac{\underline{d}_1^3+\bar{d}_5^3}{3}+5b\underline{d}_1^2+3b\bar{d}_5^2-8b^2(\underline{d}_1+\bar{d}_5)+\frac{32b^3}{3} & ,\text{if } b \ge \underline{d}_1 \ge \bar{d}_5 \\ \frac{\underline{d}_1^3+\bar{d}_5^3}{3}+3b\underline{d}_1^2+5b\bar{d}_5^2-8b^2(\underline{d}_1+\bar{d}_5)+\frac{32b^3}{3} & ,\text{if } b \ge \bar{d}_5 \ge \underline{d}_1 \\ \frac{\underline{d}_1^3+\bar{d}_5^3}{3}+b\underline{d}_1^2+3b\bar{d}_5^2-8b^2\bar{d}_5+\frac{20b^3}{3} & ,\text{if } \underline{d}_1 \ge b \ge \bar{d}_5 \\ \frac{\underline{d}_1^3+\bar{d}_5^3}{3}+3b\underline{d}_1^2+b\bar{d}_5^2-8b^2\underline{d}_1+\frac{20b^3}{3} & ,\text{if } \bar{d}_5 \ge b \ge \underline{d}_1 \\ \frac{\underline{d}_1^3+\bar{d}_5^3}{3}+b\underline{d}_1^2+7b\bar{d}_5^2-16b^2\bar{d}_5+\frac{32b^3}{3} & ,\text{if } \underline{d}_1 \ge \bar{d}_5 \text{ and } \frac{4b}{3} \ge \bar{d}_5 \ge b \\ \frac{\underline{d}_1^3+\bar{d}_5^3}{3}+7b\underline{d}_1^2+b\bar{d}_5^2-16b^2\underline{d}_1+\frac{32b^3}{3} & ,\text{if } \bar{d}_5 > \underline{d}_1 \text{ and } \frac{4b}{3} \ge \underline{d}_1 \ge b \\ \frac{\underline{d}_1^3+\bar{d}_5^3}{3}+b\left(\underline{d}_1^2+\bar{d}_5^2\right) & ,\text{if } \underline{d}_1 \ge \frac{4b}{3} \text{ and } \bar{d}_5 \ge \frac{4b}{3} \end{cases}
\end{aligned}
$$

Induction step 1:. $j-1 \to j$, j even

If Ω_1 forms a 1-partition, the best experiment has the following characteristics (cf. the results of case $j-1$): Ω_2 is a 1-partition iff $2b-\underline{d}_1 \le \min\{\frac{(j-2)b}{j-3},\max\{b,\bar{d}_j\}\}$, i.e.,

$$\underline{d}_1 \ \ge\ \min\left\{b,\max\left\{\frac{(j-4)b}{j-3},2b-\bar{d}_j\right\}\right\},$$

while Ω_2 and Ω_3 are a 2-partition iff $\underline{d}_1 \le \min\{b,\max\left\{\frac{(j-4)b}{j-3},2b-\bar{d}_j\right\}\}$.

If Ω_1 and Ω_2 form a 2-partition, the optimal value of a partial bi-pooling policy on $\omega \in (\underline{\omega},\bar{\omega})$ is given by

$$\min_d \quad v^*(\underline{d}_1,d,2)+v^*(2b-d,\bar{d}_j,j-2) \quad \text{s.t.} \quad \underline{d}_1+d \ \ge\ 2b.$$

The solution of the optimization problem is

$$d^* = \begin{cases} \frac{(j-4)b}{j-3} & ,\ \text{if } \underline{d}_1 \ge \frac{(j-2)b}{j-3} \text{ and } \bar{d}_j \le \frac{(j-4)b}{j-3} \\ b & ,\ \text{if } \underline{d}_1 \ge b \text{ and } \bar{d}_j \ge b \\ \bar{d}_j & ,\ \text{if } \underline{d}_1 \ge 2b-\bar{d}_j \text{ and } b \ge \bar{d}_j \ge \frac{(j-4)b}{j-3} \\ 2b-\underline{d}_1 & ,\ \text{if } \underline{d}_1 \le \min\{\frac{(j-2)b}{j-3},\max\{b,2b-\bar{d}_j\}\} \end{cases}$$

If $\underline{d}_1 \leq \min\{\frac{(j-2)b}{j-3}, \max\{b, 2b-\bar{d}_j\}\}$, the best experiment under which Ω_1 and Ω_2 form a 2-partition is essentially one under which Ω_1 and Ω_2 form two separate 1-partitions as the solution to the minimization problem is $d^* = 2b - \underline{d}_1$. Consequently, the globally optimal experiment must be the best one under which Ω_1 is a 1-partition.

Notice that $\min\{\frac{(j-2)b}{j-3}, \max\{b, 2b-\bar{d}_j\}\} \geq \min\{b, \max\{\frac{(j-4)b}{j-3}, 2b-\bar{d}_j\}\}$. Hence, if $\underline{d}_1 \geq \min\{\frac{(j-2)b}{j-3}, \max\{b, 2b-\bar{d}_j\}\}$, implying that $\underline{d}_1 > \min\{b, \max\{\frac{(j-4)b}{j-3}, 2b-\bar{d}_j\}\}$, the best experiment under which Ω_1 forms a 1-partition also belongs to the constraint set of the minimization problem. But it is not optimal as $d^* \neq 2b - \underline{d}_1$. The globally optimal experiment is thus one under which Ω_1 belongs to a 2-partition.

One obtains

$$v^*(\underline{d}_1, \bar{d}_j, j) = \begin{cases} v^*(\underline{d}_1, \underline{d}_1, 1) + v^*(2b - \underline{d}_1, \bar{d}_j, j-1) & \text{, if } \underline{d}_1 \leq \min\{\frac{(j-2)b}{j-3}, \max\{b, \bar{d}_j\}\} \\ v^*(\underline{d}_1, d^*, 2) + v^*(\frac{(j-2)b}{j-3}, 2b - d^*, j-2) & \text{, if } \underline{d}_1 \geq \min\{\frac{(j-2)b}{j-3}, \max\{b, \bar{d}_j\}\} \end{cases}$$

By plugging in the terms of the functions $v^*(\cdot, \cdot, j-1)$ and $v^*(\cdot, \cdot, j-2)$, one obtains the expressions in the lemma's statement.

Induction step 2:. $j-1 \to j$, j odd

If Ω_1 forms a 1-partition, the best experiment has these characteristics (cf. the results of case $j-1$): Ω_2 is a 1-partition iff $2b - \underline{d}_1 \leq \min\{\frac{(j-3)b}{j-4}, \max\{b, 2b-\bar{d}_j\}\}$, i.e.,

$$\underline{d}_1 \geq \min\left\{b, \max\left\{\frac{(j-5)b}{j-4}, \bar{d}_j\right\}\right\},$$

while Ω_2 and Ω_3 are a 2-partition iff $\underline{d}_1 \geq \min\left\{b, \max\left\{\frac{(j-5)b}{j-4}, \bar{d}_j\right\}\right\}$.

If Ω_1 and Ω_2 form a 2-partition, the optimal value of a partial bi-pooling policy on $\omega \in (\underline{\omega}, \bar{\omega})$ is given by

$$\min_d v^*(\underline{d}_1, d, 2) + v^*(2b - d, \bar{d}_j, j-2) \quad \text{s.t.} \quad \underline{d}_1 + d \geq 2b.$$

The solution of the optimization problem is

$$d^* = \begin{cases} \frac{(j-3)b}{j-2} & \text{, if } \underline{d}_1 \geq \frac{(j-1)b}{j-2} \text{ and } \bar{d}_j \geq \frac{(j-1)b}{j-2} \\ b & \text{, if } \underline{d}_1 \geq b \text{ and } \bar{d}_j \leq b \\ 2b - \bar{d}_j & \text{, if } \underline{d}_1 \geq \bar{d}_j \text{ and } \frac{(j-1)b}{j-2} \geq \bar{d}_j \geq b \\ 2b - \underline{d}_1 & \text{, if } \underline{d}_1 \leq \min\{\frac{(j-1)b}{j-2}, \max\{b, \bar{d}_j\}\} \end{cases}.$$

If $\underline{d}_1 \leq \min\{\frac{(j-1)b}{j-2}, \max\{b, \bar{d}_j\}\}$, the best experiment under which Ω_1 and Ω_2 form a 2-partition is essentially one under which Ω_1 and Ω_2 form two separate 1-partitions as the solution to the minimization problem is $d^* = 2b - \underline{d}_1$. Consequently, the globally optimal experiment must be the best one under which Ω_1 is a 1-partition.

Notice that $\min\{\frac{(j-1)b}{j-2}, \max\{b, \bar{d}_j\}\} \geq \min\{b, \max\{\frac{(j-5)b}{j-4}, \bar{d}_j\}\}$. As a consequence, if $\underline{d}_1 \geq \min\{\frac{(j-1)b}{j-2}, \max\{b, \bar{d}_j\}\}$, implying that $\underline{d}_1 \geq \min\{b, \max\{\frac{(j-5)b}{j-4}, \bar{d}_j\}\}$, the best experiment under which Ω_1 forms a 1-partition also belongs to the constraint set of the minimization problem. But it is not optimal as $d^* \neq 2b - \underline{d}_1$. The globally optimal experiment is thus one under which Ω_1 belongs to a 2-partition.

One obtains

$$
\begin{aligned}
&v^*(\underline{d}_1, \bar{d}_j, j) \\
&= \begin{cases} v^*(\underline{d}_1, \underline{d}_1, 1) + v^*(2b - \underline{d}_1, \bar{d}_j, j-1) & \text{, if } \underline{d}_1 \leq \min\{\frac{(j-1)b}{j-2}, \max\{b, \bar{d}_j\}\} \\ v^*(\underline{d}_1, d^*, 2) + v^*(\frac{(j-2)b}{j-3}, 2b - d^*, j-2) & \text{, if } \underline{d}_1 \geq \min\{\frac{(j-1)b}{j-2}, \max\{b, \bar{d}_j\}\}\} \end{cases}
\end{aligned}
$$

By plugging in the terms of the functions $v^*(\cdot, \cdot, j-1)$ and $v^*(\cdot, \cdot, j-2)$, one obtains the expressions in the lemma's statement. □

Optimal Partial Bi-pooling Policy Given $\underline{d}_1 + \bar{d}_j$. Next, I determine the optimal partial bi-pooling policy for any fixed $\underline{d}_1 + \bar{d}_j \in (0, 2b]$ by comparing the optimal partial bi-pooling policies for fixed $\underline{d}_1$ and fixed $\bar{d}_j$ from above. It turns out that it is symmetric in the sense that $\underline{d}_1 = \bar{d}_j$:

Lemma 1.51. *For any $j \in \mathbb{N}$, and any $\underline{d}_1, \bar{d}_j \in (0, 2b)$ so that $\underline{d}_1 + \bar{d}_j \in (0, 2b]$, it holds that $v^*(\underline{d}_1, \bar{d}_j, j) \geq v^*\left(\frac{\underline{d}_1 + \bar{d}_j}{2}, \frac{\underline{d}_1 + \bar{d}_j}{2}, j\right)$.*

Proof. This can be easily checked by inspection of v^* as specified above. □

Optimal Bi-Pooling Policies. The aim of this section is to determine the optimal bi-pooling policy for a given bias $b \in \left[\frac{1}{2n}, \frac{1}{2(n-1)}\right)$, where $n \geq 2$.

By Fact 3, all incentive compatibility constraints are binding under the optimal bi-pooling policy of size n. Notice that $\underline{d}_1 + (n-1)2b + \bar{d}_n = 1$, which yields $\underline{d}_1 + \bar{d}_n = 1 - (n-1)2b \in \left(0, \frac{1}{n}\right] \subset (0, 2b)$. Therefore, the results on optimal partial bi-pooling policies also apply to the complete bi-pooling policies on the whole state space: Set $\underline{\omega} = 0$ and $\bar{\omega} = 1$.

It can be easily verified that the bi-pooling policies of size n as characterized in Lemma 1.40 and Lemma 1.39 are optimal because their payoffs correspond to v^*.

The payoff of the optimal bi-pooling policy of size n is weakly decreasing in b on the interval $\left[\frac{1}{2n}, \frac{1}{2(n-1)}\right)$. This is obvious because the IC constraints become stricter as b increases, so the value of the minimization problem weakly increases. The payoff of the uniform monotone partition of size $n-1$ is constant in b. So for given $b \in \left[\frac{1}{2n}, \frac{1}{2(n-1)}\right)$, the globally optimal experiment is the best bi-pooling policy of size n if b is sufficiently large, while it is the uniform monotone partition of size $n-1$ if b is sufficiently small.

Proof of Proposition 1.42. Fix $b \in \left[\frac{1}{2n}, \frac{1}{2(n-1)}\right)$. Note that $\mathscr{U}_R^{BP} = 0$. Moreover, notice that $\mathscr{U}_R^* \geq -\frac{1}{12(n-1)^2} \geq -\frac{b^2}{3(1-2b)^2}$, where the first inequality follows from the fact that the payoff is at least the one achieved by a uniform monotone partition of size $n-1$, and the second inequality follows from $b \geq \frac{1}{2n}$. Finally, notice that $\mathscr{U}_R'' \leq -\frac{1}{12N} - \frac{b^2(N^2-1)}{3}$, where $2bN(N-1) = 1$ (cf. Crawford and Sobel (1982)). Substituting N, it can be easily verified that $\mathscr{U}_R'' - 2\mathscr{U}_R^* \leq -\frac{1}{12N} - \frac{b^2(N^2-1)}{3} + \frac{2b^2}{3(1-2b)^2} \leq 0$ for any $b \leq 0.19$. If $b > 0.19$, $\mathscr{U}_R^* \geq -\frac{1}{48}$ and $\mathscr{U}_R'' \leq \frac{-49}{600}$, hence $\mathscr{U}_R'' - 2\mathscr{U}_R^* < 0$.

Proof of Proposition 1.43. For any $b \leq \frac{1}{2}$, the receiver's ex-ante expected payoff under any mediation rule is bounded above by $-\frac{1}{3}b(1-b)$ (see Goltsman et al. (2009)). By Lemma 1.38, the receiver's ex-ante expected payoff in the unique best equilibrium under public learning is bounded below by $-\frac{1}{12(n-1)^2}$, where n is the unique integer such that $b \in \left(\frac{1}{2n}, \frac{1}{2(n-1)}\right]$. It can be easily verified that this lower bound exceeds the above upper bound.

References

Aliprantis, Charalambos, and Kim Border. 2006. *Infinite Dimensional Analysis.* Springer, Berlin. [15]

Ambrus, Attila, Eduardo M Azevedo, and Yuichiro Kamada. 2013. "Hierarchical cheap talk." *Theoretical Economics* 8 (1): 233–61. [16]

Antić, Nemanja, and Nicola Persico. 2023. "Equilibrium selection through forward induction in cheap talk games." *Games and Economic Behavior.* [4, 17]

Argenziano, Rossella, Sergei Severinov, and Francesco Squintani. 2016. "Strategic information acquisition and transmission." *American Economic Journal: Microeconomics* 8 (3): 119–55. [10]

Arieli, Itai, Yakov Babichenko, Rann Smorodinsky, and Takuro Yamashita. 2023. "Optimal persuasion via bi-pooling." *Theoretical Economics* 18 (1): 15–36. [6, 9, 26, 28, 47]

Aumann, Robert J, and Sergiu Hart. 2003. "Long cheap talk." *Econometrica* 71 (6): 1619–60. [7, 34]

Blackwell, David. 1953. "Equivalent comparisons of experiments." *Annals of Mathematical Statistics* 24 (2): 265–72. [4]

Blume, Andreas, Oliver J Board, and Kohei Kawamura. 2007. "Noisy talk." *Theoretical Economics* 2 (4): 395–440. [7, 35]

Candogan, Ozan, and Philipp Strack. 2022. "Optimal disclosure of information to privately informed agents." *Theoretical Economics.* [9, 25]

Caplin, Andrew, and Mark Dean. 2013. "Behavioral implications of rational inattention with shannon entropy." Working paper. National Bureau of Economic Research. [20]

Caplin, Andrew, Mark Dean, and John Leahy. 2019. "Rational inattention, optimal consideration sets, and stochastic choice." *Review of Economic Studies* 86 (3): 1061–94. [20]

Chen, Ying, Navin Kartik, and Joel Sobel. 2008. "Selecting cheap-talk equilibria." *Econometrica* 76 (1): 117–36. [4]

Crawford, Vincent P., and Joel Sobel. 1982. "Strategic information transmission." *Econometrica* 50 (6): 1431–51. [3–7, 9, 19, 25, 29, 33, 34, 62]

Crémer, Jacques, and Richard P McLean. 1988. "Full extraction of the surplus in Bayesian and dominant strategy auctions." *Econometrica,* 1247–57. [6]

Deimen, Inga, and Dezső Szalay. 2019. "Delegated expertise, authority, and communication." *American Economic Review* 109 (4): 1349–74. [10]

Dessein, Wouter. 2002. "Authority and communication in organizations." *Review of Economic Studies* 69 (4): 811–38. [35]

Duffy, John, and Frank Heinemann. 2021. "Central bank reputation, cheap talk and transparency as substitutes for commitment: Experimental evidence." *Journal of Monetary Economics* 117: 887–903. [3]

Dworczak, Piotr, and Giorgio Martini. 2019. "The simple economics of optimal persuasion." *Journal of Political Economy* 127 (5): 1993–2048. [9]

Ely, Jeffrey C. 2017. "Beeps." *American Economic Review* 107 (1): 31–53. [9]

Gentzkow, Matthew, and Emir Kamenica. 2014. "Costly persuasion." *American Economic Review* 104 (5): 457–62. [9, 16]

Gentzkow, Matthew, and Emir Kamenica. 2016. "A Rothschild-Stiglitz approach to Bayesian persuasion." *American Economic Review* 106 (5): 597–601. [9, 25]

Goltsman, Maria, Johannes Hörner, Gregory Pavlov, and Francesco Squintani. 2009. "Mediation, arbitration and negotiation." *Journal of Economic Theory* 144 (4): 1397–420. [62]

Green, Jerry R, and Nancy L Stokey. 2007. “A two-person game of information transmission.” *Journal of Economic Theory* 135 (1): 90–104. [5]

Ivanov, Maxim. 2010. “Informational control and organizational design.” *Journal of Economic Theory* 145 (2): 721–51. [7, 10, 25, 33, 35]

Kamenica, Emir, and Matthew Gentzkow. 2011. “Bayesian persuasion.” *American Economic Review* 101 (6): 2590–615. [4, 8, 9, 16, 34]

Kleiner, Andreas, Benny Moldovanu, and Philipp Strack. 2021. “Extreme points and majorization: Economic applications.” *Econometrica* 89 (4): 1557–93. [9, 28, 47]

Kolotilin, Anton, Tymofiy Mylovanov, Andriy Zapechelnyuk, and Ming Li. 2017. “Persuasion of a privately informed receiver.” *Econometrica* 85 (6): 1949–64. [25]

Krishna, Vijay, and John Morgan. 2004. “The art of conversation: eliciting information from experts through multi-stage communication.” *Journal of Economic Theory* 117 (2): 147–79. [7, 25, 34]

Lipnowski, Elliot. 2020. “Equivalence of Cheap Talk and Bayesian Persuasion in a Finite Continuous Model.” Note. [6]

Lipnowski, Elliot, and Doron Ravid. 2020. “Cheap talk with transparent motives.” *Econometrica* 88 (4): 1631–60. [9]

Lopomo, Giuseppe, Luca Rigotti, and Chris Shannon. 2022. “Detectability, duality, and surplus extraction.” *Journal of Economic Theory* 204: 105465. [6, 19]

Lou, Yichuan. 2023. “Sender-Optimal Learning and Credible Communication.” Working paper. [9]

Lyu, Qianjun, and Wing Suen. 2022. “Information design in cheap talk.” Working paper. [9]

Maćkowiak, Bartosz, Filip Matějka, and Mirko Wiederholt. 2018. “Survey: Rational inattention, a disciplined behavioral model.” Working paper. [6]

Matějka, Filip, and Alisdair McKay. 2015. “Rational inattention to discrete choices: A new foundation for the multinomial logit model.” *American Economic Review* 105 (1): 272–98. [6]

Matysková, Ludmila, and Alfonso Montes. 2021. “Bayesian Persuasion With Costly Information Acquisition.” Working paper. [9]

Pei, Harry D. 2015. “Communication with endogenous information acquisition.” *Journal of Economic Theory* 160: 132–49. [7, 10, 11, 25, 33, 35]

Ravid, Doron, Anne-Katrin Roesler, and Balázs Szentes. 2022. “Learning before trading: on the inefficiency of ignoring free information.” *Journal of Political Economy* 130 (2): 346–87. [26]

Yang, Ming. 2015. “Coordination with flexible information acquisition.” *Journal of Economic Theory* 158: 721–38. [4]

Chapter 2

On-The-Match Search and Match-Specific Productivity Growth*

Joint with Axel Niemeyer and Finn Schmieter

2.1 Introduction

We study decentralized match formation with search frictions as in Shimer and Smith (2000) and Smith (2006) with two new features: agents continue to search even when they are matched (on-the-match search), and matches become more productive the longer they persist (match-specific productivity growth). The combination of these two features is novel and captures many interesting scenarios. For example, workers are looking for better jobs, firms are trying to replace underperforming workers, businesses are looking for better contractors, and romantic relationships sometimes fall apart as one partner meets somebody else.[1] In many of these situations, partnerships become more productive the longer they persist, as agents become attuned to one another and learn how to optimize their joint pro-

* We thank Daniel Bird, Ege Destan, Francesc Dilmé, Thomas Kohler, Daniel Krähmer, Stephan Lauermann, Benny Moldovanu, Justus Preusser, Tobias Rachidi, Dezsö Szalay, and seminar participants from Bonn and Cologne for valuable comments and suggestions. Kreutzkamp and Niemeyer gratefully acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2126/1 – 390838866. Kreutzkamp and Schmieter gratefully acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project B04).

1. In the labor market, on-the-job search by workers is a well-documented empirical fact (see Fallick and Fleischman (2004)), but even on the firm side, there is evidence for replacement hiring (see Burgess, Lane, and Stevens (2000) and Albak and Sørensen (1998)). Bobbio (2019) and Acharya and Wee (2020) argue that replacement hiring can partly be explained through "on-the-job" search by firms for better workers. This is particularly prominent in the labor market for CEOs and managers (Parrino (1997); Murphy and Zabojnik (2007)). For evidence regarding the marriage market, see Stevenson and Wolfers (2007).

duction over time.[2] We are interested in the matching behavior and sorting patterns that emerge with on-the-match search and how they are affected by match-specific productivity growth.

In our model, there are two types of agents, more productive agents (H-agents) and less productive agents (L-agents), and a continuum of agents for each type. Time is continuous, and meetings between agents are Poisson events. If two agents meet, they observe each others' types. If they agree to match, they each enjoy a flow utility that depends on their own as well as their partner's type. Single agents of each type enter the market at a constant rate, and agents leave the market at random times. What we have described so far is our baseline model, where agents search on-the-match but flow utility in a match is constant over time. In an augmented model, we additionally assume that flow utility is an increasing function of match tenure, reflecting match-specific productivity growth. Our solution concept is steady state equilibrium: we require that agents' matching behavior is mutually optimal given the distribution of matched and unmatched agents and that this distribution is in a steady state.[3]

Let us now summarize our main findings regarding the baseline model. Without on-the-match search, assortative matching is typically driven by the agents' search for more productive partners—by the market clearing top-down—but with on-the-match search, the behavior of less productive agents becomes much more relevant to the sorting patterns that emerge in equilibrium.[4] The reason is that agents trade-off flow utility against the stability of a match, i.e., the rate at which their partner leaves for another agent. Indeed, in our baseline model, less productive L-agents may prefer to match among themselves for fear of being left by the more productive H-agents; thus, the market may additionally clear bottom-up. The fact that agents' acceptance strategies need not be monotone—that it may be more difficult for more productive agents to find a partner—contrasts with many other matching models where agents do not search on-the-match.

Our baseline model has a fundamental problem with equilibrium multiplicity, which is rooted in the fact that match stability is endogenous: stability both determines and is determined by the agents' behavior. For example, if all other L-agents accept to match with H-agents rather than L-agents, then an LL-match becomes less stable, making H-agents more desirable for any individual L-agent. The same kind of

2. For empirical evidence, see the literature on team-specific human capital (Kellogg (2011); Jaravel, Petkova, and Bell (2018); Chen (2021)). In the context of employment, this is also known as firm-specific human capital (see Becker (1964); Topel (1991); Lazear (2009)).

3. We keep our model as simple and tractable as possible, e.g., there are only two types of agents and utility is non-transferable, while retaining the distinguishing features of the search-and-matching framework as well as the solution concept of Shimer and Smith (2000) and Smith (2006). Our model applies to both one-sided markets and two-sided markets with symmetric sides.

4. With transferable utility, one additionally needs type complementaries in production for assortative matching to obtain. See Chade, Eeckhout, and Smith (2017, Section 4) for a comprehensive survey of sorting without on-the-match search.

coordination problem also governs the behavior of H-agents. Specifically, in an HH-match, it is optimal to stay in the match if the partner plans to stay, but it can also be optimal to leave for an L-agent if the partner plans to leave, precisely because becoming single entails a period without any flow utility. The fact that matched agents play a kind of coordination game with respect to staying in or leaving their current match implies a continuum of steady state equilibria in our baseline model. Strikingly, in almost every equilibrium, matched agents miscoordinate on payoff-dominated behavior.

We find that match-specific productivity growth can alleviate the problem of equilibrium multiplicity: if flow utility increases in the duration of a match, then miscoordination can no longer be sustained in equilibrium. In a nutshell, the reason is that agents anticipate their match to become stable far into the future, breaking miscoordination as an equilibrium even slightly earlier, and this argument can be iterated backward in time. Consequently, by letting productivity growth vanish, we can select precisely those steady state equilibria of the baseline model where agents coordinate on payoff-dominant behavior—we call them coordination equilibria. In a coordination equilibrium, H-agents always accept to match with H-agents, agents never accept to match with agents of their partner's type, and matched L-agents either accept to match with L-agents or H-agents but not with both. We fully characterize the set of coordination equilibria. Specifically, there are at most three equilibria, there is at most one pure-strategy equilibrium where matched L-agents accept to match with L-agents (the "assorting equilibrium"), and there is at most one pure-strategy equilibrium where matched L-agents accept to match with H-agents (the "non-assorting equilibrium"); we provide conditions for equilibrium uniqueness or coexistence.

More generally, match-specific productivity growth can significantly affect agents' matching behavior, which makes it an important factor in assessing the performance of decentralized matching markets. Compared to our baseline model, two new and opposing effects are at play. On the one hand, a growing flow utility stabilizes any type of match in the long run, even those between H-agents and L-agents, leading to less assortative matching.[5] On the other hand, a growing flow utility also incentivizes agents to foster growth in matches that are stable to begin with: becoming single now not only hurts because it entails a period without production but also because it means forgoing productivity accumulation in another, more stable match. This incentivizes L-agents to match among themselves, leading to more assortative matching. Either effect can dominate in equilibrium depending on the precise shape

5. We assume that productivity growth bites in that agents lock into any kind of match once enough time has passed, but we make no assumptions about how long a match must persist for agents to change their behavior. This "lock-in" effect is consistent with evidence from labor markets that job-to-job transitions decrease with employment tenure (see Hall (1982)).

of productivity growth, even to the extent that the predictions of the baseline model may be reversed.

The rest of the paper is organized as follows. In Sections 2.2 and 2.3, we introduce our baseline model with on-the-match search and define the notion of steady state equilibrium. In Sections 2.4 and 2.5, we analyze the baseline model, first the set of coordination equilibria and then all the other equilibria where agents fail to coordinate. In Section 2.6, we augment our baseline model with match-specific productivity growth. In Sections 2.7 and 2.8, we analyze the augmented model, first characterizing only partial equilibria and then taking into account the steady state balance conditions. In Section 2.9, we present an argument for selecting coordination equilibria in the baseline model through vanishing rates of match-specific productivity growth. In Section 2.10, we are then in a position to compare the two models and assess the effects of match-specific productivity growth. Section 3.5 concludes. All proofs are in the Appendix. The remainder of this introduction discusses the related literature and thus clarifies our contribution.

Related Literature. As mentioned before, we build on the search-and-matching framework with heterogeneous agents of Shimer and Smith (2000) and Smith (2006). To this framework, we add two new features: on-the-match search and match-specific productivity growth. As far as we are aware, our paper is the first to examine these two features jointly.

On-the-match search is related to on-the-job search, which search-theoretic models of the labor market have long taken into account (see the survey by Rogerson, Shimer, and Wright (2005).) However, there is a crucial difference between the two notions. In a typical labor market model with on-the-job search, firms create jobs through vacancy posting, and once a vacancy is filled, a firm cannot fire or replace its worker while the worker can search for other open vacancies. Translated to our setting, this means that only one agent in a match searches for better partners and only so among the unmatched, i.e., firms do not search "on-the-job". Therefore, individual decision problems are quite different: an agent whose partner does not search on-the-match no longer needs to worry about stability, which is a key determinant of matching behavior and sorting patterns when everyone searches on-the-match. For this reason, search-and-matching models with heterogeneous workers, heterogeneous firms, and on-the-job search are different from our baseline model (see, e.g., Dolado, Jansen, and Jimeno (2009), Gautier, Teulings, and Vuuren (2010), and Lentz (2010)). Bobbio (2019) and Acharya and Wee (2020) consider on-the-job search by firms in models with heterogeneous or ex ante identical agents, respectively, but their scope and methodology are very different from ours.

Only few other papers study on-the-match search in the sense described above. Cornelius (2003) is close to our baseline model in that she also assumes two types of agents and non-transferable utility, but there are three key assumptions she makes that we do not. First, agents that enter a stable match—a match that will not be dissolved in equilibrium—immediately leave the market and are replaced by two identical singles.[6] Second, matches between two more productive H-agents are stable, and agents never leave their partner for an agent of the partner's type. Third, agents break indifferences in favor of accepting a match. We show that many of the insights in Cornelius (2003) obtain in our model with *exogenous* entry and exit, in particular, that there is at most one non-assorting and at most one assorting equilibrium ("Type I" and "Type II" equilibria, respectively, in Cornelius (2003)) and therefore, that matching behavior may be a matter of coordination on what is stable. Moreover, we give a selection argument based on-match-specific productivity growth that rules out the miscoordination that Cornelius (2003) assumes away from the start. Finally, we show that tie-breaking without cloning is problematic because pure-strategy equilibria need no longer exist. Nevertheless, Cornelius (2003) examines two features from which we abstract, namely different search intensities in-and-out of a match and divorce costs.

The idea that stability can be a matter of coordination is also explored in Burdett, Imai, and Wright (2004). In their model, ex-ante identical agents choose their search intensity in a match, hence might search more intensely if their partner does the same.

There are two more recent papers on on-the-match search: Bartolucci and Monzon (2019) and Goldmanis and Ray (2019). Bartolucci and Monzon (2019) study a model with finitely many types and transferable utility. They show that with on-the-match search, assortative matching can emerge even without complementaries in production when there are high enough search frictions. Goldmanis and Ray (2019) study a two-sided market with a continuum of types on each side and non-transferable utility. They give conditions for the existence of an equilibrium in which agents upgrade to higher types of partners whenever possible, akin to our non-assorting equilibrium, and show that the associated sorting pattern converges to perfect assortative matching as agents become infinitely-lived.[7] In short, these models are quite sophisticated, ours is simpler; they study the implications of on-the-match search on sorting in specific situations, we study these implications more generally.

Finally, match-specific productivity growth has been considered in the context of labor markets by Pissarides (1994) and Schwartz (2020), but the models, methodology, and scope of these papers are again completely different from ours.

6. This "cloning" assumption is arguably problematic because it neglects the impact that agents' matching decisions have on the pool of unmatched agents; see Chade, Eeckhout, and Smith (2017, Section 4) for a discussion.

7. Without on-the-match search and non-transferable utility, a series of papers has pointed out the phenomenon of "block segregation", a sorting pattern that is not perfectly assortative; see again Chade, Eeckhout, and Smith (2017, Section 4) and in particular Smith (2006) for a discussion. In our setting with binary types, block segregation is moot.

2.2 Baseline Model: On-The-Match Search

Time is continuous. At any given moment, there is a continuum of agents in the market, each of whom is either single or matched to another agent. Each agent is characterized by a permanent productivity type, which we assume to be either L or H (low or high) for simplicity.[8] We call an agent of type $i \in \{L, H\}$ an *i-agent*, and we call a match between an i-agent and a j-agent an *ij-match*.

New i-agents enter the market as singles at a constant rate η_i, i.e., in an interval of time dt, a mass $\eta_i dt$ of new single i-agents enters the market. Agents exit the market according to a Poisson process with rate δ; thus, the lifetime of each agent is an exponential random variable. If a matched agent exits the market, her partner becomes single.[9]

Search is undirected and time-consuming but otherwise costless. For concreteness, we assume that meetings between agents follow a quadratic search technology (see Diamond and Maskin (1979)): every agent (matched or not) randomly meets other agents at a Poisson rate λm, where m is the total mass of agents in the market, and λ is a velocity parameter that captures the underlying degree of search frictions. When two agents meet, they observe each others' types and then decide whether or not to accept the match. If both accept, their current matches (if they are matched) are dissolved, and their former partners become single. The assumption of on-the-match search, that even matched agents continue to meet others, is what distinguishes our baseline model from much of the existing literature.[10]

Without loss of generality, we normalize the flow utility of single agents to 0. When an i-agent is matched with a j-agent, she obtains a flow utility $u_{ij} > 0$. We assume that flow utility is non-transferable across partners.[11] To ascribe meaning

8. Our baseline model is embedded in any model of on-the-match search with more than two types of agents. Therefore, our analysis remains relevant for the general case.

9. Instead of assuming market entry and exit, we could alternatively assume infinitely-lived agents and random match destruction as in Shimer and Smith (2000) and Smith (2006) without changing the results. The advantage of our approach is that we do not need to assume any discounting beyond the exit rate δ; see below.

10. Note that the search velocity λ is independent of whether an agent is single or matched. This assumption simplifies our analysis because it will be a dominant strategy for singles to accept any match. If on-the-match search is more efficient, e.g., because one has access to a larger network of potential mates, then matching behavior will remain the same. On the other hand, if off-the-match search is more efficient, e.g., because one has less time to search, then it can be optimal for singles to reject a match in order to search more efficiently. Therefore, our results might change, but the underlying trade-offs we characterize are still present. As such, equal search efficiencies for single and matched agents are a useful abstraction.

11. Non-transferable utility is a reasonable assumption in the marriage market or in labor markets where wages are fixed prior to a meeting, e.g., when employment contracts are non-negotiable as in the public sector. From a technical perspective, allowing for transfers comes at the expense of tractability: we would have to extend our model in that agents who meet now bargain over the joint surplus created by their match. Nonetheless, the underlying trade-off between flow utility and stability remains, and we would thus expect similar qualitative results when utility is transferable (and match values are supermodular in types).

to types, we assume that H-partners yield a higher flow utility than L-partners, i.e., $u_{iL} < u_{iH}$ for $i \in \{L, H\}$.[12] Note that the exit rate δ implicitly acts as a discount rate on future flow utility as agents no longer gain utility when they exit the market. For simplicity, we assume no further discounting beyond the exit rate.

Although we follow much of the related literature in framing our model as a one-sided rather than a two-sided market, i.e., a market where agents can be divided into two groups and a match consists of exactly one agent from each group, the gist of our findings carries over to two-sided markets. Indeed, our one-sided model is equivalent to a symmetric two-sided model where agents' strategies and inflows of single agents are the same on both market sides.

2.3 The Definition of Steady State Equilibrium

The goal of this section is to define an appropriate notion of steady state equilibrium for our baseline model with on-the-match search. Steady state equilibrium imposes two requirements on the matching behavior and the distribution of agents over the six possible matching states, i.e., L-agents or H-agents being single, matched with an L-agent, or matched with an H-agent. First, the agents' matching behavior is mutually optimal given the match distribution (partial equilibrium). Second, this distribution is stable over time given the agents' matching behavior (steady state). In the following, we shall formalize these two requirements one at a time.

2.3.1 Partial Equilibrium

We need some preliminary definitions. A *stationary strategy* for an i-agent is a tuple $p_i = (p_{ijk})_{j,k \in \{L,H\}}$, where $p_{ijk} \in [0, 1]$ is the probability that an i-agent matched with a j-agent accepts to match with a k-agent. In our model, single agents will accept to match with everyone they meet, as this yields a positive flow utility and meeting opportunities continue to arrive at the same rate.[13] Therefore, we shall spare ourselves the hassle of carrying along the strategies of singles and take their behavior as given throughout. Moreover, we restrict attention to type-symmetric strategy profiles in which agents of the same type use the same strategy.

Suppose agents take as given the distribution of matches, i.e., a tuple of masses $\boldsymbol{m} = (m_{L\emptyset}, m_{H\emptyset}, m_{LL}, m_{LH}, m_{HL}, m_{HH})$, where m_{ij} is the mass of i-agents that are matched with j-agents and $m_{i\emptyset}$ is the mass of single i-agents. Clearly, the account-

12. Whether or not flow utility is increasing in one's own type is irrelevant because types are permanent, but this would certainly be a reasonable assumption.

13. One could imagine that agents make their strategies dependent on the matching state or even matching history of the agents they meet. In this case, there could be additional equilibria in which single agents do not accept to match with everyone they meet because they would otherwise be discriminated against in future interactions. We abstract from such considerations.

ing identity $m_{LH} = m_{HL}$ holds, and we shall sometimes identify these two masses without further mention.

Suppose agents follow strategies $\boldsymbol{p} = (p_L, p_H)$. Let $\alpha_{ik}(\boldsymbol{m},\boldsymbol{p})$ denote the rate at which an i-agent meets a k-agent that is willing to match with her, and let $\beta_{ij}(\boldsymbol{m},\boldsymbol{p})$ denote the rate at which an i-agent that is matched with a j-agent becomes single because her partner leaves the market or rematches. Formally, we have

$$\alpha_{ik}(\boldsymbol{m},\boldsymbol{p}) = \lambda\left(m_{k\emptyset} + \sum_{l\in\{L,H\}} p_{kli} m_{kl}\right) \tag{2.1}$$

$$\beta_{ij}(\boldsymbol{m},\boldsymbol{p}) = \delta + \sum_{k\in\{L,H\}} p_{jik}\alpha_{jk}(\boldsymbol{m},\boldsymbol{p}). \tag{2.2}$$

Equipped with the definition of these rates, we can recursively define the expected continuation utility $V_{ij}(\boldsymbol{m},\boldsymbol{p})$ of an i-agent that is matched with a j-agent (or single, in which case we write $j = \emptyset$) for given masses $\boldsymbol{m}$ and strategies $\boldsymbol{p}$. For matched agents, we have

$$V_{ij}(\boldsymbol{m},\boldsymbol{p}) = \frac{u_{ij} + \sum_{k\in\{L,H\}} p_{ijk}\alpha_{ik}(\boldsymbol{m},\boldsymbol{p})V_{ik}(\boldsymbol{m},\boldsymbol{p}) + \beta_{ij}(\boldsymbol{m},\boldsymbol{p})V_{i\emptyset}(\boldsymbol{m},\boldsymbol{p})}{\delta + \sum_{k\in\{L,H\}} p_{ijk}\alpha_{ik}(\boldsymbol{m},\boldsymbol{p}) + \beta_{ij}(\boldsymbol{m},\boldsymbol{p})}. \tag{2.3}$$

Let us parse this expression. The denominator is the rate at which an ij-match is dissolved, which happens either if the i-agent exits the market, or if she rematches with another agent, or if she becomes single. Equivalently, the reciprocal of the denominator is the expected match duration. Now, consider the three terms in the numerator separately, each divided by the denominator. The first term is the expected utility gained from the current match; the second term is the probability of transitioning to an ik-match when the current match is dissolved times the respective expected continuation utility; the third term is the probability of becoming single when the current match is dissolved, again, times the expected continuation utility. The expression $V_{i\emptyset}(\boldsymbol{m},\boldsymbol{p})$ for single agents ($j = \emptyset$) is similar:

$$V_{i\emptyset}(\boldsymbol{m},\boldsymbol{p}) = \frac{\sum_{k\in\{L,H\}} \alpha_{ik}(\boldsymbol{m},\boldsymbol{p})V_{ik}(\boldsymbol{m},\boldsymbol{p})}{\delta + \sum_{k\in\{L,H\}} \alpha_{ik}(\boldsymbol{m},\boldsymbol{p})}. \tag{2.4}$$

Definition 2.1. The tuple $(\boldsymbol{m},\boldsymbol{p})$ is a *partial equilibrium* if $V_{ij}(\boldsymbol{m},\boldsymbol{p}) > V_{ik}(\boldsymbol{m},\boldsymbol{p})$ implies $p_{ijk} = 0$ and $V_{ij}(\boldsymbol{m},\boldsymbol{p}) < V_{ik}(\boldsymbol{m},\boldsymbol{p})$ implies $p_{ijk} = 1$.

Our definition requires that agents' acceptance decisions are optimal under the condition of being accepted. It thus rules out equilibria in weakly dominated strategies where i-agents never accept j-agents because j-agents never accept i-agents.[14]

We conclude the discussion of partial equilibrium behavior by characterizing the agents' fundamental trade-off between flow utility and stability.

Lemma 2.2. *The following are equivalent.*

(1) $V_{ij}(\boldsymbol{m},\boldsymbol{p}) > V_{ik}(\boldsymbol{m},\boldsymbol{p})$.

(2)

$$\frac{u_{ij}}{u_{ik}} > \frac{\delta + \sum_{l\in\{L,H\}} \alpha_{il}(\boldsymbol{m},\boldsymbol{p}) + \beta_{ij}(\boldsymbol{m},\boldsymbol{p})}{\delta + \sum_{l\in\{L,H\}} \alpha_{il}(\boldsymbol{m},\boldsymbol{p}) + \beta_{ik}(\boldsymbol{m},\boldsymbol{p})}. \tag{2.5}$$

Proof. See Appendix 2.A.1. □

In plain words, an i-agent matched with a k-agent accepts to match with a j-agent if and only if the relative gain in flow utilities (left-hand side) exceeds the relative increase in the costs of instability (right-hand side), i.e., the rate $\beta_{ij}(\boldsymbol{m},\boldsymbol{p})$ at which i is left by her potential new partner versus the rate $\beta_{ik}(\boldsymbol{m},\boldsymbol{p})$ at which she is left by her current partner. Note the following comparative statics: if the exit rate δ or the rate $\sum_{k\in\{L,H\}} \alpha_{ik}(\boldsymbol{m},\boldsymbol{p})$ of finding a partner when being single are high, then the right-hand side is close to 1. In either case, stability has no value because the agent expects to leave the market soon or expects to be single for only a little while. Thus, she makes her matching decisions almost exclusively based on flow payoffs. Conversely, if these rates are low, then the rate of being left by the partner dominates the right-hand side, hence stability becomes an essential consideration.

We say that a match is *stable* given $(\boldsymbol{m},\boldsymbol{p})$ if it dissolves only due to market exit, i.e., at rate 2δ.

2.3.2 Steady States

Given agents' strategies $\boldsymbol{p} = (p_L, p_H)$, the market is in a steady state if, for each of the six possible matching states, the inflow and outflow of agents exactly balance. Formally, for all $i,j \in \{L,H\}$, the masses $\boldsymbol{m}$ must satisfy the *balance conditions*

$$\eta_i + \sum_{k\in\{L,H\}} m_{ik}\beta_{ik}(\boldsymbol{m},\boldsymbol{p}) = m_{i\emptyset}\left(\delta + \sum_{k\in\{L,H\}} \alpha_{ik}(\boldsymbol{m},\boldsymbol{p})\right) \tag{2.6}$$

$$\alpha_{ij}(\boldsymbol{m},\boldsymbol{p})\left(m_{i\emptyset} + \sum_{k\in\{L,H\}} m_{ik}p_{ikj}\right) = m_{ij}\big(\beta_{ij}(\boldsymbol{m},\boldsymbol{p}) + \beta_{ji}(\boldsymbol{m},\boldsymbol{p})\big), \tag{2.7}$$

14. By the one-shot deviation principle, the above definition of partial equilibrium implies the standard definition that each agent's strategy should maximize her expected lifetime utility given the strategies of others.

where the first equation is for singles, the second equation is for matched agents, and for both equations, the inflow is on the left-hand side, whereas the outflow is on the right-hand side.[15] For example, L-agents enter the pool of singles at the exogenous entry rate η_i or when their current partner leaves them, and they leave the pool of singles when they exit the market (at Poisson rate δ) or when they meet someone that is willing to match with them.

Definition 2.3. The tuple $(\boldsymbol{m}, \boldsymbol{p})$ is a *steady state* if the masses $\boldsymbol{m}$ satisfy the balance conditions (2.6) and (2.7) for the given strategy profile $\boldsymbol{p}$.

Definition 2.4. The tuple $(\boldsymbol{m}, \boldsymbol{p})$ is a *steady state equilibrium* if it is both a partial equilibrium and a steady state.

2.4 Coordination Equilibria

We divide the analysis of steady state equilibria into two parts. In this section, we analyze a particular class of steady state equilibria, which we call coordination equilibria. In a coordination equilibrium, H-agents always accept to match with other H-agents, and only when they are single do they accept to match with L-agents, agents never accept to match with agents of the same type as their current partner, and matched L-agents either accept to match with other L-agents or with H-agents but not with both. In the next section, we analyze the remaining steady state equilibria of the baseline model.

Definition 2.5. A steady state equilibrium $(\boldsymbol{m}, \boldsymbol{p})$ is a *coordination equilibrium* if
(1) $p_{HHL} = 0$ and $p_{HLH} = 1$,
(2) $p_{ijj} = 0$ for all $i, j \in \{L, H\}$, and
(3) $p_{LHL} = 0$ or $p_{LLH} = 0$.

Each restriction corresponds to a situation where agents coordinate on payoff-dominant behavior. In an HH-match, if both agents stay together until one exits the market, then they both enjoy the highest possible expected continuation utility across every pair of their strategies, regardless of how all other agents in the market behave. However, if one H-agent is concerned that her H-partner might leave for an L-agent, then it might be best for her to match with the next L-agent she meets so as not to become single and thus unproductive later on. This coordination problem

15. These balance conditions are exactly the same for symmetric two-sided markets. Indeed, in one-sided markets with discrete types, as considered in this paper, there is the subtle issue that if two agents from the same matching state meet and agree to match, then *both* exit the current state and *both* enter the same new state. Thus, the respective entry and exit rates are twice the respective match formation and dissolution rates. However, under the quadratic meeting technology, the aggregate meeting rate of agents in the same matching state is only $1/2\lambda m_{ij}^2$, where m_{ij} is the respective mass of such agents; thus, the balance conditions for one-sided and symmetric two-sided markets remain the same.

bears some resemblance to a stag hunt game. Similarly, in any ij-match, if neither agent leaves for an agent of their partner's type, then both agents are strictly better off because their match becomes more stable, yet they gain no additional flow utility by switching. Finally, if L-agents switch back and forth between H-agents and other L-agents, then there is excessive match destruction so that any such steady state equilibrium would be Pareto-dominated by one where, all other strategies equal, L-agents make only the net switches, thus switch only in one direction.

In a coordination equilibrium, the behavior of H-agents is fully determined, but it remains to characterize the behavior of L-agents. We distinguish two cases. In an assorting equilibrium, L-agents match with L-agents unless they are matched to one already. In a non-assorting equilibrium, L-agents match with H-agents instead.

Definition 2.6. A coordination equilibrium $(\boldsymbol{m}, \boldsymbol{p})$ is

- *assorting* if $p_{LHL} = 1$ and $p_{LLH} = 0$,
- *non-assorting* if $p_{LHL} = 0$ and $p_{LLH} = 1$,
- *mixed assorting* if $0 \leq p_{LHL} < 1$ and $p_{LLH} = 0$, and
- *mixed non-assorting* if $p_{LHL} = 0$ and $0 \leq p_{LLH} < 1$.[16]

The following theorem is about the existence and uniqueness of such equilibria.

Theorem 2.7. *There exists a coordination equilibrium, and there is at most one of each of the equilibria in Theorem 2.6. Moreover, keeping all other parameters fixed, there are cutoff values $\bar{u}, \underline{u}, \hat{u} > 1$, where $\hat{u} > \max\{\bar{u}, \underline{u}\}$, such that*

- *the assorting equilibrium exists if and only if $\frac{u_{LH}}{u_{LL}} \leq \bar{u}$,*
- *the non-assorting equilibrium exists if and only if $\frac{u_{LH}}{u_{LL}} \geq \underline{u}$,*
- *the mixed assorting equilibrium exists if and only if $\bar{u} < \frac{u_{LH}}{u_{LL}} \leq \hat{u}$,*
- *the mixed non-assorting equilibrium exists if and only if $\underline{u} < \frac{u_{LH}}{u_{LL}} \leq \hat{u}$,*
- *the two mixed equilibria coincide if and only if $\frac{u_{LH}}{u_{LL}} = \hat{u}$.*

Proof. See Section 2.A.2. □

The result is illustrated in Figure 2.1. It says that the existence of the different types of coordination equilibria is determined by the ratio $\frac{u_{LH}}{u_{LL}}$, i.e., the relative gain in flow utility of an L-agent in a match with an H-agent rather than an L-agent. Specifically, the assorting equilibrium is the unique coordination equilibrium whenever the relative gain in flow utility from an LH-match is sufficiently small so that, in view of Theorem 2.2, the stability of an LL-match is more important. Similarly, the non-assorting equilibrium is the unique coordination equilibrium whenever the gain in flow utility is sufficiently large so that match stability is less important. In the intermediate range of flow payoffs, both of these equilibria may coexist ($\underline{u} < \bar{u}$)

16. An equilibrium where $p_{LHL} = 0$ and $p_{LLH} = 0$, i.e., where there is no on-the-match search by L-agents, is clearly not mixed. However, it is helpful think of this knife-edge case as the coincidence of a mixed assorting and a mixed non-assorting equilibrium.

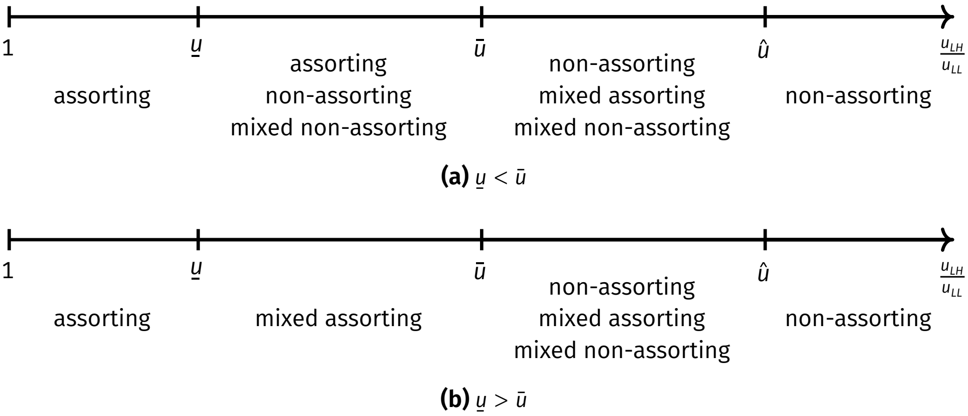


Figure 2.1. Existence of coordination equilibria as a function of $\frac{u_{LH}}{u_{LL}}$.

or neither equilibrium may exist at all ($\underline{u} > \bar{u}$), in which case there are only mixed coordination equilibria that make an L-agent exactly indifferent between H-agents and L-agents. Indeed, one can find parameter constellations such that $\underline{u} < \bar{u}$, and one can find other parameter constellations such that $\underline{u} > \bar{u}$.

Why is it that the two pure coordination equilibria—assorting and non-assorting—may coexist or not exist at all? The answer lies in the way the two equilibria differ, namely the behavior of L-agents and the steady state masses. On the one hand, the behavior of L-agents determines whether or not an LL-match is stable: if agents behave as in the assorting equilibrium, then the match is stable, but if they behave as in the non-assorting equilibrium, it is not. Thus, there is an incentive for L-agents to follow the behavior of other L-agents, and it is precisely this collective coordination among L-agents that allows the two equilibria to coexist. On the other hand, different partial equilibrium behavior also implies different steady state masses. In particular, if agents behave as in the assorting equilibrium, then there are relatively many single H-agents. Thus, there is an incentive for an L-agent to act against the other L-agents by matching with an H-agent because she can expect to quickly find a new partner when her H-partner leaves. When we increase the ratio $\frac{u_{LH}}{u_{LL}}$ towards the intermediate range of values, this steady state effect can be more pronounced than the coordination effect described above; in other words, matching with L-agents can stop being an equilibrium before matching with H-agents starts being an equilibrium, leaving mixing as the only possibility.

Let us describe in more detail how the steady state masses differ across the various coordination equilibria. Indeed, the comparative statics are exactly as one might expect; the steady state matching is more assortative in the assorting equilibrium than in the non-assorting equilibrium.

Theorem 2.8. *When comparing coordination equilibria, if p_{LHL} is higher and p_{LLH} is lower, then there are fewer single L-agents, more single H-agents, fewer LH-matches,* *and more LL-matches. The mass of HH-matches is the same in every coordination equilibrium.*

Proof. See Section 2.A.2. □

The results of this section show that with on-the-match search, sorting is driven not only by the more productive H-agents seeking to match among themselves but also by the behavior of the less productive L-agents. The fact that an L-agent may accept to match with another L-agent even when matched with an H-agent contrasts with many other matching models where acceptance strategies are monotone, where higher types are accepted more often than lower types and the market clears top-down (see also the discussion in Section 2.10). In our model, it is the type preferred by the L-agents that determines how the market clears, whether we observe more or less assortative matching. Indeed, if multiple coordination equilibria coexist, then the sorting pattern is a matter of coordination among L-agents.

2.5 Miscoordination and Equilibrium Multiplicity

Equilibrium multiplicity is a fundamental problem when agents search on-the-match, and our baseline model is no exception. The reason is the following: if agents believe that their match is stable and therefore accept other matches less often, then this belief becomes a self-fulfilling prophecy, and the match is actually more stable. However, this reasoning could just as easily be reversed, yielding multiple equilibria. Indeed, in the previous section, we have already seen how collective coordination among L-agents can yield two equilibria, the assorting and the non-assorting one. Both equilibria are plausible in that they cannot be attributed to coordination failure. In this section, we present two results on the existence of equilibria where agents clearly fail to coordinate, even on a bilateral level. One result is about H-agents matching with L-agents rather than H-agents, and the other is about agents matching with agents of their partner's type. Therefore, it is not without loss of generality to consider only the coordination equilibria from the previous section.

Theorem 2.9. *Keeping all other parameters fixed, there is a cutoff value* $\tilde{u} > 1$ *such that there exists a steady state equilibrium* $(\boldsymbol{m}, \boldsymbol{p})$ *where*
(1) $p_{HHL} = 1$ *and* $p_{HLH} = 0$*, and*
(2) $p_{ijj} = 0$ *for all* $i, j \in \{L, H\}$
if and only if $\frac{u_{HH}}{u_{HL}} \leq \tilde{u}$.

Proof. See Appendix 2.A.3. □

The intuition for this result is again based on the trade-off between flow utility and stability (Lemma 2.2): if H-agents expect to be abandoned by other H-agents, then they might seek stability in an HL-match and thereby forgo the higher flow utility in an HH-match. The condition for existence in Theorem 2.9—that H-agents are not significantly more productive with H-agents than with L-agents—can be satisfied independently of the conditions in Theorem 2.7. Thus, for certain parameter constellations, there may indeed exist an assorting equilibrium, a non-assorting equilibrium, and an equilibrium with miscoordination among H-agents. However, equilibrium multiplicity is much more pronounced, as shown by the following result.

Theorem 2.10. *There exists a continuum of steady state equilibria.*

Proof. See Appendix 2.A.3. □

For this result, we construct equilibria in which agents sometimes replace their partner with an agent of the partner's type. This kind of multiplicity may seem contrived in that it can be overcome by assuming a tie-breaking rule, e.g., that agents only accept matches they strictly prefer. However, recall from Theorem 2.7 that the existence of a steady state equilibrium can only be guaranteed if agents are allowed to play mixed strategies. Thus, tie-breaking is out of question if we wish to guarantee equilibrium existence.

Theorems 2.9 and 2.10 show that our baseline model cannot sharply predict equilibrium behavior by itself; it requires an external selection argument. This raises the following question: Should we restrict attention to coordination equilibria for the sake of payoff-dominance (as in the previous section)? Not necessarily. If we recall the analogy between the agents' coordination problem and a stag hunt game, there are certainly reasons why one might expect miscoordination among agents, perhaps most notably because coordination is a matter of trust, and mistrust is pervasive in the real world. Nevertheless, we will later provide an argument that rationalizes coordination through match-specific productivity growth (see Section 2.9). For this, however, we must first extend our baseline model.

2.6 Augmented Model: Match-Specific Productivity Growth

In this section, we augment our baseline model with match-specific productivity growth. Formally, the flow utility $u_{ij}(t) > 0$ of an i-agent matched with a j-agent is now a strictly increasing function of match duration t, whereas it was constant over time before. In particular, accumulated productivity is lost once a match is dissolved, and agents do not become more productive on their own.[17] This extension will also require some adjustments in the definition of steady state equilibrium.

17. In the context of labor markets, the literature distinguishes between general human capital and specific human capital that is tied to a particular firm or occupation (see e.g. Becker (1964)). The type of productivity growth we analyze is analogous to the accumulation of specific human capital, but we abstract from general productivity growth, e.g., agents randomly transitioning between types, to keep our model simple. However, our qualitative insights into matching behavior and sorting would remain valid even with the addition of general productivity growth because agents would still have to make a trade-off between flow utility and stability.

We make three assumptions about match-specific productivity growth. First, the utility from any given match is finite, which also ensures that all expected continuation utilities are finite: for all $i, j \in \{L, H\}$,

$$\int_0^\infty u_{ij}(t) e^{-2\delta t} \, dt \; < \; \infty. \tag{2.8}$$

Second, as before, H-agents are more productive partners than L-agents, and we now assume this pointwise for every match duration: $u_{iL}(t) < u_{iH}(t)$ for all $i \in \{L, H\}$ and $t \geq 0$. Third, productivity in a match eventually increases to the point where matched agents no longer accept to match with others. This assumption makes it interesting to study match-specific productivity growth because it causes agents to change their matching behavior relative to the baseline model. Indeed, the simplest way to ensure that matched agents eventually become unwilling to match with others is to assume unbounded productivity growth: for all $i, j \in \{L, H\}$,

$$\lim_{t \to \infty} u_{ij}(t) \; = \; \infty. \tag{2.9}$$

Given these assumptions, our baseline model is not nested in the augmented model.

It might be more realistic to assume productivity growth that diminishes over time and converges to a (large enough) finite limit. For our analysis, however, it is not the precise limit that is important but only the question of whether agents eventually stop accepting other matches. Thus, we assume unbounded growth for simplicity. Indeed, our assumption is mild regardless of whether growth is bounded or unbounded because we allow for arbitrarily slow growth in either case.

Given match-specific productivity growth, we need to make some adjustments to the definition of steady state equilibrium. A *stationary strategy* for an i-agent is a now a collection of measurable functions $p_{ijk} : [0, \infty) \to [0, 1]$, where $p_{ijk}(t)$ is the probability that an i-agent who is matched with a j-agent for time t accepts to match with a k-agent. As before, single agents accept to match with everyone they meet.

Let $m_{ij}(t)$ denote the mass density of i-agents that have been matched with j-agents for exactly time t; thus $m_{ij} = \int_0^\infty m_{ij}(t)$. Let $\boldsymbol{m}$ denote the tuple of the single masses $m_{i\emptyset}$ and the mass densities $m_{ij}(t)$ of matched agents. For the rate α_{ij} at which an i-agent meets a j-agent that is willing to match and the rate β_{ij} at which an i-agent is left by a j-agent, we now have

$$\alpha_{ik}(\boldsymbol{m}, \boldsymbol{p}) \; = \; \lambda \left(m_{k\emptyset} + \sum_{l \in \{L,H\}} \int_0^\infty p_{kli}(t) m_{kl}(t) \, dt \right) \tag{2.10}$$

$$\beta_{ij}(\boldsymbol{m}, \boldsymbol{p}; t) \; = \; \delta + \sum_{k \in \{L,H\}} p_{jik}(t) \alpha_{jk}(\boldsymbol{m}, \boldsymbol{p}). \tag{2.11}$$

Note that β_{ij} is now a function of match duration since the same holds for the strategies $p_{jik}(t)$.

In order to define expected continuation utilities, we first define the survival function $q_{ij}(\boldsymbol{m},\boldsymbol{p};t)$ of an ij-match, i.e., the probability that an ij-match lasts for at least time t:

$$q_{ij}(\boldsymbol{m},\boldsymbol{p};t) \;=\; \exp\left(-\int_0^t \left(\beta_{ij}(\boldsymbol{m},\boldsymbol{p};s) + \beta_{ji}(\boldsymbol{m},\boldsymbol{p};s)\right) ds\right). \tag{2.12}$$

For the expected continuation utility $V_{ij}(\boldsymbol{m},\boldsymbol{p};t)$ of an i-agent that is matched with a j-agent for time t, we then have (dropping the dependence on $(\boldsymbol{m},\boldsymbol{p})$ for readability)

$$\begin{aligned} V_{ij}(t) \;=\; &\int_t^\infty \frac{q_{ij}(s)}{q_{ij}(t)} u_{ij}(s)\; ds + \int_t^\infty \frac{q_{ij}(s)}{q_{ij}(t)} \sum_{k\in\{L,H\}} \alpha_{ik} p_{ijk}(s) V_{ik}(0)\; ds \\ &+ \int_t^\infty \frac{q_{ij}(s)}{q_{ij}(t)} \beta_{ij}(s) V_{i\emptyset}\; ds. \end{aligned} \tag{2.13}$$

The three components have the same interpretation as in the baseline model, namely the expected utility from the current match, the expected continuation utility from transitioning to another match, and the expected continuation utility from becoming single. Indeed, if productivity does not grow and agents play time-invariant strategies, then (2.13) reduces to (2.3) from before. The expression $V_{i\emptyset}$ for single agents is exactly (2.4) from earlier and, in particular, time-invariant.

The definition of partial equilibrium is the same as in the baseline model, modulo the dependence on match duration, i.e., $V_{ij}(\boldsymbol{m},\boldsymbol{p};t) > V_{ik}(\boldsymbol{m},\boldsymbol{p};0)$ implies $p_{ijk}(t) = 0$ and $V_{ij}(\boldsymbol{m},\boldsymbol{p};t) < V_{ik}(\boldsymbol{m},\boldsymbol{p};0)$ implies $p_{ijk}(t) = 1$.

The definition of steady states is more involved than in the baseline model. We must now not only ensure that the six masses m_{ij} are constant over time but also that the entire mass densities $m_{ij}(t)$ with respect to match duration are constant over time. However, the following section shows that partial equilibrium behavior takes a certain form, which later allows us to simplify the balance conditions considerably.

2.7 Cutoff Strategies and Coordination

We now show that with match-specific productivity growth, agents behave as in a coordination equilibrium except that matches become stable for long enough match durations. In particular, agents use cutoff strategies with respect to match duration.

Theorem 2.11. *If $(\boldsymbol{m},\boldsymbol{p})$ is a partial equilibrium, then for all $t > 0$ and $i,j \in \{L,H\}$, $p_{HHL}(t) = 0$ and $p_{ijj}(t) = 0$. Moreover, there exist non-negative cutoffs $\boldsymbol{t} = (t_{LLH}, t_{LHL}, t_{HLH})$ such that*

$$p_{ijk}(t) = \begin{cases} 1 & \text{if } t < t_{ijk}, \\ 0 & \text{if } t > t_{ijk} \end{cases}$$

with the following properties:
(1) $t_{LLH} = 0$ *or* $t_{LHL} = 0$*,*
(2) $t_{LHL} < t_{HLH}$*, and*
(3) for each cutoff t_{ijk}*, if* $t_{ijk} > 0$*, then* t_{ijk} *is the unique solution to* $V_{ij}(t_{ijk}) = V_{ik}(0)$*.*

Proof. See Appendix 2.B.1. □

Why is miscoordination no longer an equilibrium when productivity in a match grows over time? At first glance, it would seem that there could be equilibria in which agents do not separate until a certain point in time but then suddenly begin to accept other matches because they expect their partner to do the same. However, we show by a kind of "unraveling argument" that such behavior cannot occur in equilibrium. First, we observe that agents strictly prefer to stay in a match for long enough match durations regardless of whether their partner accepts other matches since the flow utility eventually grows large enough to compensate for any lack of stability. More so, agents anticipate this "lock-in" point and therefore reject other matches even slightly earlier because it is unlikely that their partner will meet someone in the short timeframe until the match becomes stable. Thus, the match is actually stable earlier, and this logic can be iterated backward in time. Indeed, in an HH-match, this unraveling halts only at the time of matching, and for the other types of matches, we eventually reach a cutoff point.

Remark 3. The cutoff t_{HLH} is the same across all partial equilibria because it solves

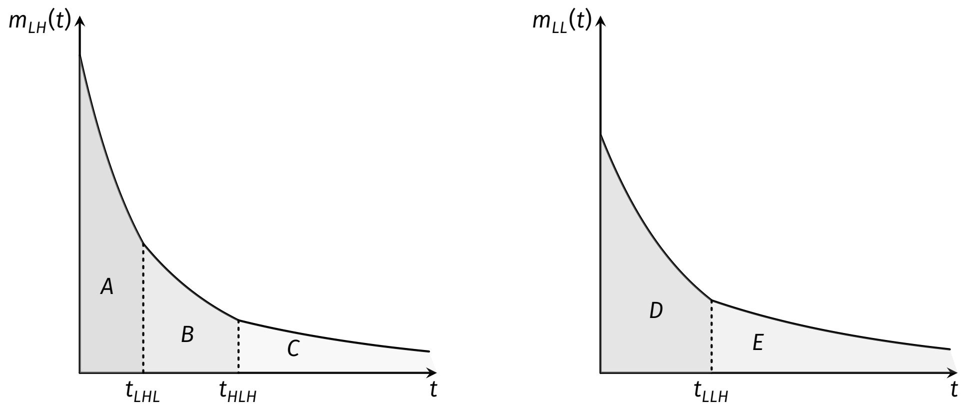
$$V_{HL}(t_{HLH}) = V_{HH}(0) \iff \int_{t_{HLH}}^{\infty} u_{HL}(t) e^{-2\delta(t - t_{HLH})} dt = \int_0^{\infty} u_{HH}(t) e^{-2\delta t} dt. \tag{2.14}$$

By Theorem 2.11, we may identify partial equilibrium strategies $\boldsymbol{p}$ with their associated cutoffs $\boldsymbol{t}$, and we thus sometimes write $(\boldsymbol{m}, \boldsymbol{t})$ for a partial equilibrium $(\boldsymbol{m}, \boldsymbol{p})$.[18]

2.8 Steady State Analysis

Our characterization of partial equilibrium behavior greatly simplifies the balance conditions that must be satisfied in a steady state, allowing for a tractable analysis of steady state equilibria.

18. Note that we have only characterized partial equilibrium behavior up to a finite number of match durations, namely $t = 0$ and the cutoffs. We can henceforth ignore this issue since matching decisions on a nullset of match durations do not affect the steady state masses.



(a) The mass density $m_{LH}(t)$ decreases exponentially over time; the rate changes at the cutoffs t_{LHL} and t_{HLH}.

(b) The mass density $m_{LL}(t)$ decreases exponentially over time; the rate changes at the cutoff t_{LLH}.

Figure 2.2. An illustration of the mass densities $m_{LH}(t)$ and $m_{LL}(t)$ of LH-matches and LL-matches, respectively, which are proportional to the survival functions $q_{LH}(t)$ and $q_{LL}(t)$.

Let us recall from the previous section that agents use cutoff strategies in every partial equilibrium $(\boldsymbol{m},\boldsymbol{p})$. These cutoff strategies imply that a match can only be in one of a few states, depending on the match duration. Specifically, there are only three states for an LH-match: (1) if the agents have been matched for a time shorter than t_{LHL}, then both accept to match with other agents, and therefore, the match dissolves at rate $2\delta + \beta_{LH}(\boldsymbol{m},\boldsymbol{p};t) + \beta_{HL}(\boldsymbol{m},\boldsymbol{p};t)$, where $\beta_{LH}(\boldsymbol{m},\boldsymbol{p};t)$ and $\beta_{HL}(\boldsymbol{m},\boldsymbol{p};t)$ are constant over $(0,t_{LHL})$; (2) if the agents have been matched for a time longer than t_{LHL} but shorter than t_{HLH}, then only the H-agent agrees to match with other agents, and consequently, the match dissolves at rate $2\delta + \beta_{LH}(\boldsymbol{m},\boldsymbol{p};t)$, where $\beta_{LH}(\boldsymbol{m},\boldsymbol{p};t)$ is constant over (t_{LHL},t_{HLH}); (3) for durations longer than t_{HLH}, the match is stable and henceforth dissolves at rate 2δ. Similarly, an LL-match can only be in one of two states depending on whether it has lasted for longer or shorter than t_{LLH}, and for an HH-match, there is only one state since the match is stable to begin with. Hence, the survival function $q_{ij}(\boldsymbol{m},\boldsymbol{p};t)$ of an ij-match is exponentially decreasing with a different but constant rate in each state; this observation is illustrated in Figure 2.2.

If we integrate the mass density $m_{ij}(t)$ of an ij-match—which is proportional to the survival function $q_{ij}(t)$—over one of the intervals in the previous paragraph, we obtain the mass of agents in the associated state. Specifically, we define

$$A = \int_0^{t_{LHL}} m_{LH}(t)\,dt, \qquad B = \int_{t_{LHL}}^{t_{HLH}} m_{LH}(t)\,dt, \qquad C = \int_{t_{HLH}}^{\infty} m_{LH}(t)\,dt,$$
$$D = \int_0^{t_{LLH}} m_{LL}(t)\,dt, \qquad E = \int_{t_{LLH}}^{\infty} m_{LL}(t)\,dt.$$

These masses are the respective areas under the curves in Figure 2.2. By Theorem 2.11, $A = 0$ or $D = 0$. Henceforth, we write $\boldsymbol{m} = (m_{L\emptyset}, m_{H\emptyset}, A, B, C, D, E, m_{HH})$ for the tuple of all the relevant masses.

With the definition of these masses, we can formulate the balance conditions. The reason is that the time-invariance of (A, B, C) implies the time-invariance of the entire mass density $m_{LH}(t)$, and, similarly, the time-invariance of (D, E) implies the time-invariance of the entire mass density $m_{LL}(t)$. In other words, we can simplify the balance conditions from tracking entire mass densities to tracking masses in a few states. For illustration purposes, let us describe here only the balance condition for A, i.e., the mass of L-agents matched with H-agents for a time shorter than t_{LHL}, and move the other balance conditions to Appendix 2.B.2 (see equations (2.B.1) to (2.B.8)). The balance condition for A reads:

$$\begin{aligned}\lambda m_{H\emptyset}(m_{L\emptyset} + D) \;=\; & 2\delta A + \lambda A(m_{L\emptyset} + A) + \lambda A(m_{H\emptyset} + A + B) \\ & + \lambda m_{H\emptyset}(m_{L\emptyset} + D) q_{LH}(\boldsymbol{m}, \boldsymbol{p}; t_{LHL}),\end{aligned}$$

where

$$q_{LH}(\boldsymbol{m}, \boldsymbol{p}; t_{LHL}) = \exp\left(-(2\delta + \lambda(m_{L\emptyset} + A) + \lambda(m_{H\emptyset} + A + B))t_{LHL}\right).$$

The left-hand side is the inflow of L-agents into LH-matches, which is, by definition, the inflow into A. The right-hand side is the outflow from A, which comprises the outflow due to (1) market exit, (2) the L-agent rematching, (3) the H-agent rematching, and (4) the flow from A into B through matches that last for longer than t_{LHL}.

A steady state equilibrium is, as before, a partial equilibrium that satisfies the balance conditions. We have the following result about the existence and uniqueness of steady state equilibria in the augmented model.

Theorem 2.12. *There exists a steady state equilibrium. Moreover, fixing all parameters except the inflows (η_L, η_H), for every pair of single masses $(m_{L\emptyset}, m_{H\emptyset})$, there exists a unique pair of inflows (η_L, η_H), a unique collection of masses $\mathbf{m}$ with single masses $(m_{L\emptyset}, m_{H\emptyset})$, and a unique strategy profile $\mathbf{p}$ such that $(\mathbf{m}, \mathbf{p})$ is a steady state equilibrium.*

Proof. See Appendix 2.B.2. □

The fact that the balance conditions are more involved relative to the baseline model prevents us from obtaining as sharp of a uniqueness result as Theorem 2.7 for coordination equilibria in the baseline model. Nonetheless, we can show that the augmented model is identified from the masses of single agents. In other words, observing the masses of single agents is enough to uniquely determine the mass densities of matched agents, the agents' equilibrium behavior, and the inflows of agents into the market. As an illustration, one could consider the context of labor markets. With data on the unemployed population and suitable estimates of the parameters of our model, one could uniquely predict matching behavior and sorting in the market. Our proof strategy is similar to other results in the literature: we show that a particular composite mapping reflecting best-response and steady state dynamics is well-defined and continuous and argue that a fixed point of this mapping corresponds to a steady state equilibrium.

Remark 4. While we cannot uniquely predict equilibrium behavior in general, let us briefly discuss the two polar cases where the velocity parameter λ is very small or very large, i.e., if the search technology is very efficient or very inefficient. Indeed, fix all parameters except for λ. It can easily be shown that there exists some $\bar{\lambda} > 0$ such that for all $\lambda < \bar{\lambda}$, we have $t^*_{LLH} > 0$ in any steady state equilibrium, i.e., L-agents prefer to match with H-agents rather than L-agents. Moreover, as $\lambda \to \infty, m_{HH} \to \frac{\eta_H}{\delta}$, $m_{LL} \to \frac{\eta_L}{\delta}$, and $m_{L\emptyset}, m_{H\emptyset}, m_{LH} \to 0$ uniformly across all steady state equilibria. Analogous statements hold for the baseline model.

2.9 Equilibrium Selection in the Baseline Model

In this section, we take up the idea of selecting steady state equilibria in the baseline model by perturbing the model with small rates of match-specific productivity growth. We show that coordination equilibria are precisely those equilibria that survive such a perturbation.

Since the baseline model is not nested in the augmented model (see Section 2.6), we must first discuss how to map between the baseline model and the perturbed model. To this end, we call the limit of a sequence of steady state equilibria in the augmented model along a sequence of vanishing growth rates a limit equilibrium, and we define when such a limit equilibrium corresponds to a steady state equilibrium in the baseline model. As we will be studying the baseline and augmented model jointly, we assume that the common parameters of both models are the same and fixed throughout this section. Moreover, we fix flow utilities in the baseline model as $\boldsymbol{u}^* = \left(u^*_{ij}\right)_{i,j\in\{L,H\}}$, leaving only the flow utilities in the augmented model undetermined.

Definition 2.13. The tuple $(\boldsymbol{m},\boldsymbol{p})$ is a *limit equilibrium* of the augmented model if there exist sequences

$$\left(\boldsymbol{m}^n, \boldsymbol{p}^n\right)_{n\in\mathbb{N}} \quad \text{and} \quad \boldsymbol{u}^n = \left(\left(u^n_{ij}(t)\right)_{i,j\in\{L,H\},\, t\geq 0}\right)_{n\in\mathbb{N}}$$

of masses, strategies, and flow utilities in the augmented model with the following properties:
(1) $\lim_{n\to\infty} u^n_{ij}(t) = u^*_{ij}$ for all $i,j \in \{L,H\}$ and $t \geq 0$,
(2) $(\boldsymbol{m}^n, \boldsymbol{p}^n)$ is a steady state equilibrium when flow utilities are given by $\boldsymbol{u}^n$, and
(3) $(\boldsymbol{m}^n, \boldsymbol{p}^n) \to (\boldsymbol{m}, \boldsymbol{p})$ pointwise as $n \to \infty$.

Definition 2.14. A steady state equilibrium $(\boldsymbol{m}^*, \boldsymbol{p}^*)$ in the baseline model *corresponds* to a limit equilibrium $(\boldsymbol{m}, \boldsymbol{p})$ of the augmented model if, for all $i, j, k \in \{L, H\}$, $m^*_{i\emptyset} = m_{i\emptyset}$, $m^*_{ij} = m_{ij}$, and

$$m^*_{ij} p^*_{ijk} = \int_0^\infty m_{ij}(t) p_{ijk}(t)\, dt. \tag{2.15}$$

The second definition says that the two equilibria correspond to one another if they have the same steady state masses and the same masses of agents that accept to match with others. Instead of the latter condition, it might seem more straightforward to require that for every match duration t, the acceptance probabilities p^*_{ijk} and $p_{ijk}(t)$ in the baseline and the perturbed model, respectively, are the same. However, this alternative definition would not be particularly useful because equilibrium strategies in the augmented model are (almost everywhere) pure while they may well be mixed in the baseline model.

Theorem 2.15. *A steady state equilibrium in the baseline model corresponds to a limit equilibrium of the augmented model if and only if it is a coordination equilibrium.*

Proof. See Appendix 2.C.1. □

This result is important because it can alleviate the problem of equilibrium multiplicity that is inherent to our baseline model without productivity growth. While we have previously argued that one might expect agents to play a coordination equilibrium for the sake of payoff dominance, we have also noted that coordination failure cannot be ruled out in equilibrium, leaving ample room for equilibrium multiplicity; see Section 2.5. In contrast, the present result uniquely selects the class of coordination equilibria in the baseline model, a class of steady state equilibria for which we have a sharp characterization; see Theorem 2.7.

2.10 The Effects of Match-Specific Productivity Growth

We now discuss how match-specific productivity growth changes the predictions of the baseline model. Indeed, it is only with the selection result from the previous section that we can meaningfully compare equilibrium behavior in the two models because our characterization of coordination equilibria is sharp, whereas virtually any equilibrium behavior is possible when agents miscoordinate. We find that productivity growth can both impede and promote sorting and even reverse the matching behavior of L-agents relative to the baseline model.

Similar to the baseline model, we say that a steady state equilibrium is *assorting* if $t_{LHL} > 0$ and $t_{LLH} = 0$ and *non-assorting* if $t_{LHL} = 0$ and $t_{LLH} > 0$.

Theorem 2.16. *Fix all parameters except for the flow utilities. If* $\int_0^\infty \big(u_{HH}(t) - u_{HL}(t)\big) e^{-2\delta t}\, dt$ *is sufficiently large and* $\int_0^\infty \big(u_{LH}(t) - u_{LL}(t)\big) e^{-2\delta t}\, dt$ *is sufficiently small, then every steady state equilibrium is assorting. Moreover, if* $\int_0^\infty \big(u_{HH}(t) - u_{HL}(t)\big) e^{-2\delta t}\, dt$ *is sufficiently small or* $\int_0^\infty \big(u_{LH}(t) - u_{LL}(t)\big) e^{-2\delta t}\, dt$ *is sufficiently large, then every steady state equilibrium is non-assorting.*

Proof. See Appendix 2.C.2. □

Theorem 2.16 provides conditions on flow utilities such that the matching behavior of L-agents is the same in all steady state equilibria of the augmented model. For comparison with the baseline model, let us take the initial flow utilities $u_{ij}(0)$ in the augmented model to be the flow utilities u_{ij} from the baseline model. Then, the conditions on the initial flow utilities in Theorem 2.7—our characterization of coordination equilibria in the baseline model—can be satisfied independently of the conditions on the evolution of flow utilities in Theorem 2.16. Thus, for certain flow utilities, it can be unambiguously determined in both models whether L-agents match with L-agents or H-agents, and this behavior need not be the same, i.e., productivity growth may indeed reverse the matching behavior of L-agents.

There are two opposing forces that determine the precise effect of match-specific productivity growth. On the one hand, productivity growth locks agents into non-assortative matches for long enough match durations. On the other hand, the stability of an assortative match is now desirable not only because being single entails a period without production but also because being single means forgoing productivity accumulation in a stable match. The former leads to less assortative matching while the latter leads to more assortative matching, and either force may dominate in equilibrium depending on the precise shape of productivity growth.

These findings suggest that some caution is warranted when analyzing decentralized matching markets where match-specific productivity growth could play a role because one might otherwise draw false conclusions from observed sorting patterns. For example, assortative matching can be the result of low search frictions (high λ), thus driven by the fact that more productive agents quickly find one another, or it can be driven by the agents' desire to foster growth in long-term relationships with similar types. Likewise, non-assortative matching need not be an indication of high search frictions (low λ) per se, but it could be driven by the fact that agents voluntarily commit to their matches once they become sufficiently productive.

In the baseline model, we have already seen that it may be more difficult for H-agents to find a partner than for L-agents when L-agents coordinate to match among themselves. This phenomenon can be exacerbated by match-specific productivity growth because L-agents now have an additional incentive to match among themselves, namely to foster growth in stable relationships. In labor markets, for example, this finding may provide an explanation for how "overqualification" can lead to involuntary unemployment, perhaps even more so when productivity in a match grows over time.[19]

2.11 Conclusion

We have analyzed a search-and-matching model with heterogeneous agents in the spirit of Shimer and Smith (2000) and Smith (2006) with two new features: on-the-match search and match-specific productivity growth. With on-the-match search, agents must make a trade-off between flow utility and match stability, which is in turn determined by the agents' behavior. The fact that stability is endogenous implies a coordination problem among agents and, consequently, a fundamental problem with equilibrium multiplicity when agents search on-the-match. We have shown that this problem can be alleviated by match-specific productivity growth and have characterized the set of steady state equilibria that survive a perturbation with match-specific productivity growth. In general, the desire for stability can lead less productive agents to prefer matching with other less productive agents, which is an alternative explanation for assortative matching. Moreover, the behavior of less productive agents can significantly change under match-specific productivity growth, which makes it an important factor in assessing the performance of decentralized matching markets.

The following directions for future research seem promising to us:

(1) Extend our analysis to more than two types of agents and see how equilibrium selection through match-specific productivity growth plays out. Because there is now even more scope for equilibrium multiplicity, it will be a significant challenge for future research to identify a reasonable class of steady state equilibria within which further analysis can be carried out. (After all, a model that can predict virtually any behavior is not particularly useful.) Match-specific productivity growth will still be able to eliminate some equilibria, and a characterization of the surviving equilibria would be interesting. It is needless to say that other selection arguments would prove helpful as well.

(2) Endogenize match-specific productivity growth, for example, by having matched agents play an investment game. For this extension to be tractable, some other elements of our model would most likely have to be simplified.[20]

(3) Analyze pre-match agreements and match dissolution costs. The fact that it might be more difficult for more productive agents to find a partner in equilibrium suggests that they might sometimes want to commit to a match with a less productive agent. What are the aggregate effects of allowing agents to

19. See e.g. Galperin et al. (2020) for empirical evidence.

20. Lentz and Roys (2015) and Flinn, Gemici, and Laufer (2017) study labor market models with on-the-job search where firms may invest in match-specific productivity.

commit to a match unilaterally for some period of time? Relatedly, what happens if match dissolution is costly or if agents can agree to make it costly upon matching?

Appendix

For the sake of readability, we henceforth omit the arguments $\boldsymbol{m}$, $\boldsymbol{p}$, and t of the rates α_{ij} and β_{ij} and the expected continuation utilities V_{ij} whenever there is no risk of confusion.

Appendix 2.A Proofs: Baseline Model

2.A.1 Section 2.3

Proof of Lemma 2.2. Fix some $i, j \in \{L, H\}$, and let $k \neq j$ be the opposite type of j.

Plugging (2.4) into (2.3), we get the following expressions for the expected continuation utilities of an i-agent in a match with a j-agent and a k-agent, respectively:

$$\begin{aligned} V_{ij} &= \frac{u_{ij} + \sum_{l \in \{L,H\}} p_{ijl} \alpha_{il} V_{il} + \beta_{ij} \left(\frac{\sum_{l \in \{L,H\}} \alpha_{il} V_{il}}{\delta + \sum_{l \in \{L,H\}} \alpha_{il}} \right)}{\delta + \sum_{l \in \{L,H\}} p_{ijl} \alpha_{il} + \beta_{ij}} \\ V_{ik} &= \frac{u_{ik} + \sum_{l \in \{L,H\}} p_{ikl} \alpha_{il} V_{il} + \beta_{ik} \left(\frac{\sum_{l \in \{L,H\}} \alpha_{il} V_{il}}{\delta + \sum_{l \in \{L,H\}} \alpha_{il}} \right)}{\delta + \sum_{l \in \{L,H\}} p_{ikl} \alpha_{il} + \beta_{ik}}. \end{aligned}$$

From these two equations, one can construct a system of linear equations of the form

$$\begin{pmatrix} A_{jj} & A_{jk} \\ A_{kj} & A_{kk} \end{pmatrix} \cdot \begin{pmatrix} V_{ij} \\ V_{ik} \end{pmatrix} = \begin{pmatrix} \left(\delta + \sum_{l \in \{L,H\}} \alpha_{il} \right) u_{ij} \\ \left(\delta + \sum_{l \in \{L,H\}} \alpha_{il} \right) u_{ik} \end{pmatrix},$$

where

$$\begin{aligned} A_{jj} &= \delta \left(\delta + \sum_{l \in \{L,H\}} \alpha_{il} \right) + p_{ijk} \alpha_{ik} \left(\delta + \sum_{l \in \{L,H\}} \alpha_{il} \right) + \beta_{ij} \left(\delta + \alpha_{ik} \right), \\ A_{jk} &= -\alpha_{ik} \left[p_{ijk} \left(\delta + \sum_{l \in \{L,H\}} \alpha_{il} \right) + \beta_{ij} \right], \\ A_{kj} &= -\alpha_{ij} \left[p_{ikj} \left(\delta + \sum_{l \in \{L,H\}} \alpha_{il} \right) + \beta_{ik} \right], \\ A_{kk} &= \delta \left(\delta + \sum_{l \in \{L,H\}} \alpha_{il} \right) + p_{ikj} \alpha_{ij} \left(\delta + \sum_{l \in \{L,H\}} \alpha_{il} \right) + \beta_{ik} \left(\delta + \alpha_{ij} \right). \end{aligned}$$

Let $\boldsymbol{A} \equiv \begin{pmatrix} A_{jj} & A_{jk} \\ A_{kj} & A_{kk} \end{pmatrix}$. Since

$$\begin{aligned}
\det(\boldsymbol{A}) &= A_{jj}A_{kk} - A_{jk}A_{kj} \\
&= \left[\delta\left(\delta + \sum_{l\in\{L,H\}} \alpha_{il} + \beta_{ij}\right) - A_{jk}\right]\left[\delta\left(\delta + \sum_{l\in\{L,H\}} \alpha_{il} + \beta_{ik}\right) - A_{kj}\right] - A_{jk}A_{kj} \\
&= \delta^2\left(\delta + \sum_{l\in\{L,H\}} \alpha_{il} + \beta_{ij}\right)\left(\delta + \sum_{l\in\{L,H\}} \alpha_{il} + \beta_{ik}\right) \\
&\quad - \delta\left(\delta + \sum_{l\in\{L,H\}} \alpha_{il} + \beta_{ij}\right)A_{kj} - \delta\left(\delta + \sum_{l\in\{L,H\}} \alpha_{il} + \beta_{ik}\right)A_{jk} \\
&> 0,
\end{aligned}$$

it follows by Cramer's rule that the system of equations has a unique solution, namely

$$\begin{aligned}
V_{ij} &= \frac{1}{\det(\boldsymbol{A})} \cdot \det\begin{pmatrix} \left(\delta + \sum_{l\in\{L,H\}} \alpha_{il}\right)u_{ij} & A_{jk} \\ \left(\delta + \sum_{l\in\{L,H\}} \alpha_{il}\right)u_{ik} & A_{kk} \end{pmatrix} \\
V_{ik} &= \frac{1}{\det(\boldsymbol{A})} \cdot \det\begin{pmatrix} A_{jj} & \left(\delta + \sum_{l\in\{L,H\}} \alpha_{il}\right)u_{ij} \\ A_{kj} & \left(\delta + \sum_{l\in\{L,H\}} \alpha_{il}\right)u_{ik} \end{pmatrix}.
\end{aligned}$$

From this, we can infer that

$$\begin{aligned}
V_{ij} > V_{ik} \quad &\Leftrightarrow \quad u_{ij}A_{kk} - u_{ik}A_{jk} > u_{ik}A_{jj} - u_{ij}A_{kj} \\
&\Leftrightarrow \quad \frac{u_{ij}}{u_{ik}} > \frac{A_{jj} + A_{jk}}{A_{kj} + A_{kk}} = \frac{\delta + \sum_{l\in\{L,H\}} \alpha_{il} + \beta_{ij}}{\delta + \sum_{l\in\{L,H\}} \alpha_{il} + \beta_{ik}},
\end{aligned}$$

which completes the proof. □

2.A.2 Section 2.4

It is often an essential step in the analysis of steady state equilibria to prove the existence of unique steady state masses for arbitrary strategies and to show that this solution changes continuously in the strategies (see the so-called "fundamental matching lemma" in Shimer and Smith (2000)). For our baseline model, we first show the existence of steady states for arbitrary strategies and then argue that they are unique and change continuously for certain classes of strategy profiles, including those played in a coordination equilibrium. We do not show uniqueness for arbitrary strategy profiles, nor do we need it, and we would even question the validity of this claim.

As a preliminary step, let us derive the *aggregate balance conditions* by adding up the balance conditions for single agents and matched agents ((2.6) and (2.7), respectively), separately for L-agents and H-agents:

$$\eta_L = \delta\left(m_{L\emptyset} + m_{LL} + m_{LH}\right) \tag{2.A.1}$$

$$\eta_H = \delta\left(m_{H\emptyset} + m_{HH} + m_{LH}\right). \tag{2.A.2}$$

The left-hand side is the flow of i-agents that enter the market, and the right-hand side is the total flow of i-agents that exit the market.

Moreover, we shall repeatedly use the $\mathbb{R}^6$-valued mapping $F(\boldsymbol{m},\boldsymbol{p})$ that results from the balance conditions (2.6) and (2.7) by subtracting for each mass its right-hand side (outflow) from its left-hand side (inflow). In other words, F returns the excess inflows for given $(\boldsymbol{m},\boldsymbol{p})$. Also note that F is continuous in $(\boldsymbol{m},\boldsymbol{p})$ because it is a polynomial map.

Lemma 2.17. *For every strategy profile* $\boldsymbol{p}$ *there exist masses* $\boldsymbol{m}$ *satisfying the balance conditions (2.6) and (2.7).*

Proof. Fix an arbitrary strategy profile $\boldsymbol{p}$. Let $\boldsymbol{M}$ be the set of masses $\boldsymbol{m} \in \mathbb{R}^6_+$ satisfying the aggregate balance conditions (2.A.1) and (2.A.2). Note that $\boldsymbol{M}$ is non-empty, compact, and convex.

Fix any arbitrary strategy profile $\boldsymbol{p}$. Define a mapping $T : \boldsymbol{M} \to \boldsymbol{M}$ via

$$T(\boldsymbol{m}) = \boldsymbol{m} + \frac{1}{2(\delta + \lambda\frac{\eta_L+\eta_H}{\delta})}F(\boldsymbol{m},\boldsymbol{p}),$$

where the scaling factor ensures that $T(\boldsymbol{m}) \geq \boldsymbol{0}$ since the outflow from any matching state is bounded by the exit and meeting rates. Moreover, by construction, $T(\boldsymbol{m})$ satisfies the aggregate balance conditions since $\boldsymbol{m}$ does. Thus, $T(\boldsymbol{m}) \in \boldsymbol{M}$.

Finally, T is continuous since F is continuous. Thus, by Brouwer's fixed point theorem, T has a fixed point, and $\boldsymbol{m}$ is a fixed point of T if and only if $F(\boldsymbol{m},\boldsymbol{p}) = \boldsymbol{0}$, which completes the proof. □

Lemma 2.18. *Suppose* $\boldsymbol{P}$ *is a subset of strategy profiles such that, for every* $\boldsymbol{p} \in \boldsymbol{P}$*, there is a unique solution* $\boldsymbol{m}(\boldsymbol{p})$ *to the balance conditions (2.6) and (2.7). Then,* $\boldsymbol{m}(\boldsymbol{p})$ *is continuous over* $\boldsymbol{P}$*.*

Proof. By definition, $F(\boldsymbol{m}(\boldsymbol{p}),\boldsymbol{p}) = \boldsymbol{0}$. For any sequence $(\boldsymbol{p}^n) \subset \boldsymbol{P}$ with $\boldsymbol{p}^n \to \boldsymbol{p}^*$, we have

$$F\left(\lim_{n\to\infty} \boldsymbol{m}(\boldsymbol{p}^n),\boldsymbol{p}^*\right) = \lim_{n\to\infty} F\left(\boldsymbol{m}(\boldsymbol{p}^n),\boldsymbol{p}^n\right) = \lim_{n\to\infty} 0 = 0 = F\left(\boldsymbol{m}(\boldsymbol{p}^*),\boldsymbol{p}^*\right).$$

By the uniqueness of $\boldsymbol{m}(\boldsymbol{p})$, we must have $\lim_{n\to\infty} \boldsymbol{m}(\boldsymbol{p}^n) = \boldsymbol{m}(\boldsymbol{p}^*)$, which completes the proof. □

The following lemma shows the uniqueness of a steady state for a specific class of strategy profiles. This class may seem somewhat arbitrary at first, but we will be working with it later on. In any case, we are covering the strategies used in a coordination equilibrium.

Lemma 2.19. *Let $\boldsymbol{P}$ be the set of strategy profiles satisfying the restrictions of a coordination equilibrium, except that $p_{HLH} \in [0,1]$ and $p_{HLL} \in [0,1]$. For every $\boldsymbol{p} \in \boldsymbol{P}$ there exist unique masses $\boldsymbol{m}(\boldsymbol{p})$ such that $(\boldsymbol{m}(\boldsymbol{p}), \boldsymbol{p})$ is a steady state. Moreover, $\boldsymbol{m}(\boldsymbol{p})$ is continuous in $\boldsymbol{p}$ over $\boldsymbol{P}$.*

Proof. In view of Lemmata 2.17 and 2.18, it suffices to show that for every $\boldsymbol{p} \in \boldsymbol{P}$, there is at most one solution to the balance conditions (2.6) and (2.7), which simplify as follows given the restrictions above:

$$m_{HH} : \quad \lambda\left(m_{H\emptyset} + p_{HLH} m_{LH}\right)^2 = 2\delta m_{HH} \tag{2.A.3}$$

$$m_{LL} : \quad \lambda\left(m_{L\emptyset} + p_{LHL} m_{LH}\right)^2 = 2\delta m_{LL} + 2\lambda m_{LL}(p_{LLH} m_{H\emptyset} + p_{LLH} p_{HLL} m_{LH}) \tag{2.A.4}$$

$$m_{LH} : \quad \begin{aligned} &\lambda\left(m_{L\emptyset} + p_{LLH} m_{LL}\right) m_{H\emptyset} = 2\delta m_{LH} \\ &\qquad + \lambda m_{LH}\left(p_{LHL}(m_{L\emptyset} + m_{LH} p_{LHL}) + p_{HLH}(m_{H\emptyset} + p_{HLH} m_{LH})\right) \end{aligned} \tag{2.A.5}$$

$$m_{L\emptyset} : \quad \begin{aligned} &\eta_L + \delta(m_{LL} + m_{LH}) + \lambda p_{LLH} m_{LL} m_{H\emptyset} \\ &\qquad + \lambda p_{HLH} m_{LH}\left(m_{H\emptyset} + p_{HLH} m_{LH}\right) + 2\lambda m_{LL} p_{LLH} p_{HLL} m_{LH} \\ &\qquad\qquad = \delta m_{L\emptyset} + \lambda m_{L\emptyset}\left(m_{L\emptyset} + m_{H\emptyset} + p_{LHL} m_{LH}\right) \end{aligned} \tag{2.A.6}$$

$$m_{H\emptyset} : \quad \begin{aligned} &\eta_H + \delta(m_{HH} + m_{LH}) + \lambda p_{LHL} m_{LH}\left(m_{L\emptyset} + p_{LHL} m_{LH}\right) \\ &\qquad = \delta m_{H\emptyset} + \lambda m_{H\emptyset}\left(m_{L\emptyset} + m_{H\emptyset} + p_{LLH} m_{LL} + p_{HLH} m_{LH}\right). \end{aligned} \tag{2.A.7}$$

For the sake of contradiction, suppose for some $\boldsymbol{p} \in \boldsymbol{P}$ there are two distinct solutions, $\boldsymbol{m}$ and $\boldsymbol{m}'$.

We first show that $m_{LH} \neq m'_{LH}$. If $m_{LH} = m'_{LH}$, then the aggregate balance condition (2.A.2) together with (2.A.3) implies that $m_{HH} = m'_{HH}$ and $m_{H\emptyset} = m'_{H\emptyset}$. In turn, the aggregate balance condition (2.A.1) together with (2.A.4) implies that $m_{LL} = m'_{LL}$ and $m_{L\emptyset} = m'_{L\emptyset}$, contradicting that $\boldsymbol{m} \neq \boldsymbol{m}'$. Thus, without loss of generality, assume $m_{LH} < m'_{LH}$.

As in the previous paragraph, the aggregate balance condition (2.A.2) together with (2.A.3) implies that $m_{HH} \geq m'_{HH}$ and $m_{H\emptyset} > m'_{H\emptyset}$. In turn, the aggregate balance conditions (2.A.1) and (2.A.2) together with (2.A.4) imply that $m_{L\emptyset} > m'_{L\emptyset}$, and if $p_{LLH} = 0$, then $m_{LL} \geq m'_{LL}$.

Now, suppose $p_{LHL} = 0$. The inflow into $m_{H\emptyset}$, i.e., the left-hand side of (2.A.7), is larger under $\boldsymbol{m}'$ than under $\boldsymbol{m}$, whereas the outflow from $m_{H\emptyset}$, i.e., the right-hand side of (2.A.7), is larger under $\boldsymbol{m}$ than under $\boldsymbol{m}'$, which is a contradiction.

Finally, suppose $p_{LLH} = 0$. Compared to the argument in the previous paragraph, there is now an additional source of inflow into $m_{H\emptyset}$, namely $\lambda p_{LHL} m_{LH}\left(m_{L\emptyset} + p_{LHL} m_{LH}\right)$. The only way that the balance condition (2.A.7) can be satisfied for both $\boldsymbol{m}$ and $\boldsymbol{m}'$ is that this new source of inflow into $m_{H\emptyset}$ is so much larger given $\boldsymbol{m}$ than given $\boldsymbol{m}'$ that it offsets the larger outflow from $m_{H\emptyset}$ given $\boldsymbol{m}$ rather than $\boldsymbol{m}'$, as was argued in the previous paragraph. In particular,

$$p_{LHL}m_{L\emptyset}m_{LH} - p_{LHL}m'_{L\emptyset}m'_{LH} > p_{HLH}m_{H\emptyset}m_{LH} - p_{HLH}m'_{H\emptyset}m'_{LH}. \tag{2.A.8}$$

Then, however, the difference of inflows into $m_{L\emptyset}$ between $\boldsymbol{m}$ and $\boldsymbol{m}'$, i.e., the respective left-hand sides of (2.A.6), is smaller than the difference of outflows, i.e., the respective right-hand sides of (2.A.6), from $m_{L\emptyset}$ between $\boldsymbol{m}$ and $\boldsymbol{m}'$, which is a contradiction. □

We now define two classes of strategy profiles covered by Lemma 2.19. First, let $\boldsymbol{P}_{LHL}$ denote the set of strategy profiles satisfying the restrictions of a coordination equilibrium with $p_{LLH} = 0$, i.e., those of an assorting or mixed assorting equilibrium. Second, let $\boldsymbol{P}_{LLH}$ denote the set of strategy profiles satisfying the restrictions of a coordination equilibrium with $p_{LHL} = 0$, i.e., those of a non-assorting or mixed non-assorting equilibrium. Note that both sets can be linearly ordered via p_{LHL} and p_{LLH}, respectively.

The next two lemmas present comparative statics of the steady state solution $\boldsymbol{m}(\boldsymbol{p})$ with respect to strategy profiles $\boldsymbol{p} \in \boldsymbol{P}_{LHL}$ or $\boldsymbol{p} \in \boldsymbol{P}_{LLH}$.

Lemma 2.20. *Let $\boldsymbol{p}, \boldsymbol{p}' \in \boldsymbol{P}_{LHL}$ be so that $p_{LHL} < p'_{LHL}$, and let $\boldsymbol{m} = \boldsymbol{m}(\boldsymbol{p})$ and $\boldsymbol{m}' = \boldsymbol{m}(\boldsymbol{p}')$. Then, $m_{HH} = m'_{HH}$, $m_{H\emptyset} < m'_{H\emptyset}$, $m_{LH} > m'_{LH}$, $m_{L\emptyset} > m'_{L\emptyset}$ and $m_{LL} < m'_{LL}$.*

Proof. We shall repeatedly use the balance conditions (2.A.3)-(2.A.7), setting $p_{LLH} = 0$, $p_{HLH} = 1$, and $p_{HLL} = 0$. The fact that $m_{HH} = m'_{HH}$ then immediately follows from (2.A.3) and (2.A.2).

For the sake of contradiction, suppose $m_{LL} \geq m'_{LL}$. By (2.A.4), we have

$$m_{L\emptyset} + p_{LHL}m_{LH} \geq m'_{L\emptyset} + p'_{LHL}m'_{LH} > m'_{L\emptyset} + p_{LHL}m'_{LH}.$$

This inequality together with the aggregate balance condition for L-agents (2.A.1) implies $m_{LH} < m'_{LH}$ and $m_{L\emptyset} > m'_{L\emptyset}$. Moreover, the aggregate balance condition for H-agents (2.A.2) implies $m_{H\emptyset} > m'_{H\emptyset}$. Now, the left-hand side of (2.A.6) is strictly larger under $(\boldsymbol{m}, \boldsymbol{p})$ than under $(\boldsymbol{m}', \boldsymbol{p}')$, whereas the right-hand side of (2.A.6) is strictly larger under $(\boldsymbol{m}', \boldsymbol{p}')$ than under $(\boldsymbol{m}, \boldsymbol{p})$, which contradicts that both $(\boldsymbol{m}, \boldsymbol{p})$ and $(\boldsymbol{m}', \boldsymbol{p}')$ are steady states. Hence, $m_{LL} < m'_{LL}$, and, moreover,$m_{L\emptyset} + p_{LHL}m_{LH} < m'_{L\emptyset} + p'_{LHL}m'_{LH}$ by (2.A.4).

For the sake of contradiction, suppose $m_{L\emptyset} \leq m'_{L\emptyset}$. By the aggregate balance conditions (2.A.1) and (2.A.2), we have $m_{LH} > m'_{LH}$ and $m_{H\emptyset} < m'_{H\emptyset}$. As in the previous paragraph, these relationships yield a contradiction with (2.A.6). Hence, $m_{L\emptyset} > m'_{L\emptyset}$.

Finally, for the sake of contradiction, suppose $m_{LH} \leq m'_{LH}$. By the aggregate balance condition (2.A.2), we have $m_{H\emptyset} \geq m'_{H\emptyset}$. Now, the left-hand side of (2.A.5) is strictly larger under $(\boldsymbol{m}, \boldsymbol{p})$ than under $(\boldsymbol{m}', \boldsymbol{p}')$, whereas the right-hand side of (2.A.5) is strictly larger under $(\boldsymbol{m}', \boldsymbol{p}')$ than under $(\boldsymbol{m}, \boldsymbol{p})$, which contradicts that both $(\boldsymbol{m}, \boldsymbol{p})$ and $(\boldsymbol{m}', \boldsymbol{p}')$ are steady states. Hence, $m_{LH} > m'_{LH}$ and thus $m_{H\emptyset} < m'_{H\emptyset}$. □

Lemma 2.21. *Let $\boldsymbol{p}, \boldsymbol{p}' \in \boldsymbol{P}_{LLH}$ be such that $p_{LLH} < p'_{LLH}$, and let $\boldsymbol{m} = \boldsymbol{m}(\boldsymbol{p})$ and $\boldsymbol{m}' = \boldsymbol{m}(\boldsymbol{p}')$. Then, $m_{HH} = m'_{HH}$, $m_{H\emptyset} > m'_{H\emptyset}$, $m_{LH} < m'_{LH}$, $m_{L\emptyset} < m'_{L\emptyset}$ and $m_{LL} > m'_{LL}$. Moreover, $m_{H\emptyset} + m_{L\emptyset} < m'_{L\emptyset} + m'_{H\emptyset}$.*

Proof. We shall repeatedly use the balance conditions (2.A.3)-(2.A.7), setting $p_{LHL} = 0$, $p_{HLH} = 1$, and $p_{HLL} = 0$. The fact that $m_{HH} = m'_{HH}$ then immediately follows from (2.A.3) and (2.A.2).

For the sake of contradiction, suppose $m_{L\emptyset} \geq m'_{L\emptyset}$. By (2.A.4), we have $p_{LLH} m_{H\emptyset} \geq p'_{LLH} m'_{H\emptyset}$ or $m_{LL} \geq m'_{LL}$. In the former case, we get $m_{H\emptyset} \geq m'_{H\emptyset}$. In the latter case, we get $m_{LH} \leq m'_{LH}$ from the aggregate balance condition (2.A.1). Thus, by the aggregate balance condition (2.A.2), we have $m_{H\emptyset} \geq m'_{H\emptyset}$ and $m_{LH} \leq m'_{LH}$ in either case. From (2.A.5), it follows that $m_{H\emptyset} p_{LLH} m_{LL} \leq m'_{H\emptyset} p'_{LLH} m'_{LL}$. Now, (2.A.6) can only be fulfilled if all the previously stated inequalities hold with equality because the left-hand side of (2.A.6) is weakly larger under $(\boldsymbol{m}', \boldsymbol{p}')$ than under $(\boldsymbol{m}, \boldsymbol{p})$, whereas the right-hand side of (2.A.6) is weakly larger under $(\boldsymbol{m}, \boldsymbol{p})$ than under $(\boldsymbol{m}', \boldsymbol{p}')$, which would otherwise contradict that both $(\boldsymbol{m}, \boldsymbol{p})$ and $(\boldsymbol{m}', \boldsymbol{p}')$ are steady states. However, $p_{LLH} m_{H\emptyset} m_{LL} = p'_{LLH} m'_{H\emptyset} m'_{LL}$ and $m_{H\emptyset} = m'_{H\emptyset}$ imply $m_{LL} > m'_{LL}$, and $m_{LH} = m'_{LH}$ and $m_{L\emptyset} = m'_{L\emptyset}$ imply $m_{LL} = m'_{LL}$ via the aggregate balance condition (2.A.6), which is a contradiction. Hence, $m_{L\emptyset} < m'_{L\emptyset}$.

For the sake of contradiction, suppose $m_{LL} \leq m'_{LL}$. Then, $m_{LH} \geq m'_{LH}$ from the aggregate balance condition for L-agents (2.A.1). Consequently, $m_{H\emptyset} > m'_{H\emptyset}$ from (2.A.5), which contradicts the aggregate balance condition for H-agents (2.A.2). Hence, $m_{LL} > m'_{LL}$, and, as in the previous paragraph, we have $p_{LLH} m_{H\emptyset} m_{LL} < p'_{LLH} m'_{H\emptyset} m'_{LL}$ by (2.A.4).

Now, for the sake of contradiction, suppose $m_{H\emptyset} \leq m'_{H\emptyset}$ and $m_{LH} \geq m'_{LH}$. Then, the left-hand side of (2.A.5) is strictly larger under $(\boldsymbol{m}', \boldsymbol{p}')$ than under $(\boldsymbol{m}, \boldsymbol{p})$, whereas the right-hand side of (2.A.5) is weakly larger under $(\boldsymbol{m}, \boldsymbol{p})$ than under $(\boldsymbol{m}', \boldsymbol{p}')$, which contradicts that both $(\boldsymbol{m}, \boldsymbol{p})$ and $(\boldsymbol{m}', \boldsymbol{p}')$ are steady states. Hence, $m_{H\emptyset} > m_{H\emptyset}$ and $m_{LH} < m'_{LH}$ by the aggregate balance condition for H-agents (2.A.2).

Finally, summing up the balance conditions (2.A.6) and (2.A.7) for singles yields

$$\eta_L + \eta_H + \delta(m_{LL} + 2m_{LH} + m_{HH}) + \lambda m_{LH}^2 = \delta(m_{L\emptyset} + m_{H\emptyset}) + \lambda(m_{L\emptyset} + m_{H\emptyset})^2.$$

Using the aggregate balance conditions (2.A.1) and (2.A.2), we then have

$$2\eta_L + 2\eta_H + \lambda m_{LH}^2 = 2\delta(m_{L\emptyset} + m_{H\emptyset}) + \lambda(m_{L\emptyset} + m_{H\emptyset})^2.$$

Thus, $m_{L\emptyset} + m_{H\emptyset} < m'_{L\emptyset} + m'_{H\emptyset}$ since $m_{LH} < m'_{LH}$. □

Lemma 2.22. *The difference $V_{LH}(\boldsymbol{m}(\boldsymbol{p}),\boldsymbol{p}) - V_{LL}(\boldsymbol{m}(\boldsymbol{p}),\boldsymbol{p})$ is continuous and single-crossing from below over $\boldsymbol{P}_{LHL}$ and over $\boldsymbol{P}_{LLH}$.*

Proof. First, consider $\boldsymbol{p} \in \boldsymbol{P}_{LHL}$. For any tuple of masses $\boldsymbol{m}$, we have

$$\begin{aligned}
\alpha_{LL}(\boldsymbol{m},\boldsymbol{p}) &= \lambda(m_{L\emptyset} + p_{LHL} m_{LH}) \\
\alpha_{LH}(\boldsymbol{m},\boldsymbol{p}) &= \lambda m_{H\emptyset} \\
\beta_{LL}(\boldsymbol{m},\boldsymbol{p}) &= \delta \\
\beta_{LH}(\boldsymbol{m},\boldsymbol{p}) &= \delta + \lambda(m_{H\emptyset} + m_{LH}).
\end{aligned}$$

From the proof of Lemma 2.20, it follows that $\alpha_{LL}(\boldsymbol{m}(\boldsymbol{p}),\boldsymbol{p})$ and $\alpha_{LH}(\boldsymbol{m}(\boldsymbol{p}),\boldsymbol{p})$ are strictly increasing over $\boldsymbol{P}_{LHL}$, whereas $\beta_{LL}(\boldsymbol{m}(\boldsymbol{p}),\boldsymbol{p})$ and $\beta_{LH}(\boldsymbol{m}(\boldsymbol{p}),\boldsymbol{p})$ remain constant. Thus, the ratio $\frac{\delta+\alpha_{LL}+\alpha_{LH}+\beta_{LH}}{\delta+\alpha_{LL}+\alpha_{LH}+\beta_{LL}}$ is strictly decreasing over $\boldsymbol{P}_{LHL}$. Thus, by Lemma 2.2, the difference $V_{LH}(\boldsymbol{m}(\boldsymbol{p}),\boldsymbol{p}) - V_{LL}(\boldsymbol{m}(\boldsymbol{p}),\boldsymbol{p})$ is single-crossing from below over $\boldsymbol{P}_{LHL}$.

Second, consider $\boldsymbol{p} \in \boldsymbol{P}_{LLH}$. For any tuple of masses $\boldsymbol{m}$, we have

$$\begin{aligned}
\alpha_{LL}(\boldsymbol{m},\boldsymbol{p}) &= \lambda m_{L\emptyset} \\
\alpha_{LH}(\boldsymbol{m},\boldsymbol{p}) &= \lambda m_{H\emptyset} \\
\beta_{LL}(\boldsymbol{m},\boldsymbol{p}) &= \delta + \lambda p_{LLH} m_{LH} \\
\beta_{LH}(\boldsymbol{m},\boldsymbol{p}) &= \delta + \lambda(m_{H\emptyset} + m_{LH}).
\end{aligned}$$

From the proof of Lemma 2.21, it follows that $\alpha_{LL}(\boldsymbol{m}(\boldsymbol{p}),\boldsymbol{p}) + \alpha_{LH}(\boldsymbol{m}(\boldsymbol{p}),\boldsymbol{p})$ and $\beta_{LL}(\boldsymbol{m}(\boldsymbol{p}),\boldsymbol{p})$ are strictly increasing over $\boldsymbol{P}_{LLH}$, whereas $\beta_{LH}(\boldsymbol{m}(\boldsymbol{p}),\boldsymbol{p})$ remains constant. Thus, the ratio $\frac{\delta+\alpha_{LL}+\alpha_{LH}+\beta_{LH}}{\delta+\alpha_{LL}+\alpha_{LH}+\beta_{LL}}$ is strictly decreasing over $\boldsymbol{P}_{LHL}$. Thus, by Lemma 2.2, the difference $V_{LH}(\boldsymbol{m}(\boldsymbol{p}),\boldsymbol{p}) - V_{LL}(\boldsymbol{m}(\boldsymbol{p}),\boldsymbol{p})$ is single-crossing from below over $\boldsymbol{P}_{LLH}$.

For the continuity claim, note that by the proof of Lemma 2.2, the tuple of functions $(V_{L\emptyset}(\boldsymbol{m},\boldsymbol{p}), V_{LL}(\boldsymbol{m},\boldsymbol{p}), V_{LH}(\boldsymbol{m},\boldsymbol{p}))$ is the unique solution to a system of equations that is continuous in $(\boldsymbol{m},\boldsymbol{p})$. Hence, the solution is continuous in $(\boldsymbol{m},\boldsymbol{p})$ using the same argument as in the proof of Lemma 2.18. Moreover, by Lemma 2.18, $\boldsymbol{m}(\boldsymbol{p})$ is continuous in $\boldsymbol{p}$ over both $\boldsymbol{P}_{LHL}$ and $\boldsymbol{P}_{LLH}$, hence the claim. □

Proof of Theorem 2.7. Let $\boldsymbol{p} \in \boldsymbol{P}_{LHL} \cup \boldsymbol{P}_{LLH}$, and let $\boldsymbol{m}$ be an arbitrary tuple of masses. We have $\beta_{HH}(\boldsymbol{m},\boldsymbol{p}) \leq \beta_{HL}(\boldsymbol{m},\boldsymbol{p})$ and thus, by Lemma 2.2, $V_{HH}(\boldsymbol{m},\boldsymbol{p}) > V_{HL}(\boldsymbol{m},\boldsymbol{p})$ because $\frac{u_{HH}}{u_{HL}} > 1$. Hence, H-agents behave optimally in $(\boldsymbol{m},\boldsymbol{p})$.

Define three strategy profiles. First, let $\boldsymbol{p}^1 \in \boldsymbol{P}_{LHL}$ be such that $p_{LHL} = 1$, i.e., as in an assorting equilibrium. Second, let $\boldsymbol{p}^2 \in \boldsymbol{P}_{LLH}$ be such that $p_{LLH} = 1$, i.e., as in a non-assorting equilibrium. Third, let $\boldsymbol{p}^3 \in \boldsymbol{P}_{LHL} \cap \boldsymbol{P}_{LLH}$, i.e., $p_{LHL} = 0$ and $p_{LLH} = 0$.

Define the cutoffs

$$\frac{\delta + \sum_{l\in\{L,H\}} \alpha_{Ll}(\boldsymbol{m}(\boldsymbol{p}^1),\boldsymbol{p}^1) + \beta_{LH}(\boldsymbol{m}(\boldsymbol{p}^1),\boldsymbol{p}^1)}{\delta + \sum_{l\in\{L,H\}} \alpha_{Ll}(\boldsymbol{m}(\boldsymbol{p}^1),\boldsymbol{p}^1) + \beta_{LL}(\boldsymbol{m}(\boldsymbol{p}^1),\boldsymbol{p}^1)} \equiv \bar{u}$$
$$\frac{\delta + \sum_{l\in\{L,H\}} \alpha_{Ll}(\boldsymbol{m}(\boldsymbol{p}^2),\boldsymbol{p}^2) + \beta_{LH}(\boldsymbol{m}(\boldsymbol{p}^2),\boldsymbol{p}^2)}{\delta + \sum_{l\in\{L,H\}} \alpha_{Ll}(\boldsymbol{m}(\boldsymbol{p}^2),\boldsymbol{p}^2) + \beta_{LL}(\boldsymbol{m}(\boldsymbol{p}^2),\boldsymbol{p}^2)} \equiv \underline{u}$$
$$\frac{\delta + \sum_{l\in\{L,H\}} \alpha_{Ll}(\boldsymbol{m}(\boldsymbol{p}^3),\boldsymbol{p}^3) + \beta_{LH}(\boldsymbol{m}(\boldsymbol{p}^3),\boldsymbol{p}^3)}{\delta + \sum_{l\in\{L,H\}} \alpha_{Ll}(\boldsymbol{m}(\boldsymbol{p}^3),\boldsymbol{p}^3) + \beta_{LL}(\boldsymbol{m}(\boldsymbol{p}^3),\boldsymbol{p}^3)} \equiv \hat{u}.$$

Moreover, note that

$$\begin{aligned}
\sum_{l\in\{L,H\}} \alpha_{Ll}(\boldsymbol{m}(\boldsymbol{p}^3),\boldsymbol{p}^3) &= \lambda(m_{L\emptyset}+m_{H\emptyset}) &&= \sum_{l\in\{L,H\}} \alpha_{Ll}(\boldsymbol{m}(\boldsymbol{p}^2),\boldsymbol{p}^2)\\
& && < \sum_{l\in\{L,H\}} \alpha_{Ll}(\boldsymbol{m}(\boldsymbol{p}^1),\boldsymbol{p}^1)\\
\beta_{LL}(\boldsymbol{m}(\boldsymbol{p}^3),\boldsymbol{p}^3) &= \delta &&= \beta_{LL}(\boldsymbol{m}(\boldsymbol{p}^1),\boldsymbol{p}^1)\\
& && < \beta_{LL}(\boldsymbol{m}(\boldsymbol{p}^2),\boldsymbol{p}^2)\\
\beta_{LH}(\boldsymbol{m}(\boldsymbol{p}^3),\boldsymbol{p}^3) &= \delta+\lambda(m_{H\emptyset}+m_{LH}) &&= \beta_{LH}(\boldsymbol{m}(\boldsymbol{p}^2),\boldsymbol{p}^2)\\
& && = \beta_{LH}(\boldsymbol{m}(\boldsymbol{p}^1),\boldsymbol{p}^1)
\end{aligned}$$

(see the proof of Lemma 2.22 for the two rightmost expressions in each line.) Hence, $\hat{u} > \max\{\bar{u}, \underline{u}\}$.

Finally, by the definition of a partial equilibrium, we need $V_{LL} \geq V_{LH}$ in an assorting equilibrium, $V_{LH} \geq V_{LL}$ in a non-assorting equilibrium, and $V_{LL} = V_{LH}$ in every other coordination equilibrium. The claims regarding the pure coordination equilibria immediately follow from Lemma 2.2, and the claims regarding the mixed coordination equilibria follow from Lemma 2.22 by distinguishing cases based upon the comparison of $\hat{u}$ and $\frac{u_{LH}}{u_{LL}}$. □

Proof of Theorem 2.8. Immediately follows from Lemmata 2.20 and 2.21. □

2.A.3 Section 2.5

Proof of Theorem 2.9. Let $\boldsymbol{p}$ be a strategy profile where $p_{HHL} = 1$, $p_{HLH} = 0$ and $p_{ijj} = 0$ for all $i,j \in \{L,H\}$, and let $\boldsymbol{m}$ be any tuple of non-negative masses. Then,

$$\beta_{LH}(\boldsymbol{m},\boldsymbol{p}) = \delta \leq \delta + \lambda p_{LLH}(m_{H\emptyset}+m_{HL}) = \beta_{LL}(\boldsymbol{m},\boldsymbol{p}).$$

By Lemma 2.2, $V_{LH}(\boldsymbol{m},\boldsymbol{p}) > V_{LL}(\boldsymbol{m},\boldsymbol{p})$. Thus, if $(\boldsymbol{m},\boldsymbol{p})$ is a steady state equilibrium, then $p_{LLH} = 1$ and $p_{LHL} = 0$.

Now, let $\boldsymbol{p}$ be the strategy profile where $p_{HHL} = 1$, $p_{HLH} = 0$, $p_{ijj} = 0$ for all $i, j \in \{L, H\}$, $p_{LLH} = 1$, and $p_{LHL} = 0$, and let $\boldsymbol{M}$ be the set of non-negative masses satisfying the balance conditions (2.6) and (2.7) given $\boldsymbol{p}$. $\boldsymbol{M}$ is non-empty by Lemma 2.17. Moreover, $\boldsymbol{M}$ is compact since it is bounded and the preimage of the point $\mathbf{0}$ under the continuous function $F(\cdot, \boldsymbol{p})$.

Define

$$\tilde{u} \equiv \max_{\boldsymbol{m} \in \boldsymbol{M}} \frac{\delta + \sum_{l \in \{L,H\}} \alpha_{Hl}(\boldsymbol{m}, \boldsymbol{p}) + \beta_{HH}(\boldsymbol{m}, \boldsymbol{p})}{\delta + \sum_{l \in \{L,H\}} \alpha_{Hl}(\boldsymbol{m}, \boldsymbol{p}) + \beta_{HL}(\boldsymbol{m}, \boldsymbol{p})}.$$

The cutoff $\tilde{u}$ is well-defined since the term on the right-hand side is continuous in $\boldsymbol{m}$.

By Lemma 2.2, a desired steady state equilibrium exists if and only if $\frac{u_{HH}}{u_{HL}} \leq \tilde{u}$. □

Proof of Theorem 2.10. We will show that, for every $q \in [0, 1]$, there is a steady state equilibrium $(\boldsymbol{m}, \boldsymbol{p})$ such that $p_{HLL} = q$.

Fix an arbitrary $p_{HLL} = q$. Let $\boldsymbol{P}$ denote the set of strategy profiles where agents otherwise behave as in a coordination equilibrium. By Lemma 2.19, for any such strategy profile $\boldsymbol{p} \in \boldsymbol{P}$, there are unique masses $\boldsymbol{m}(\boldsymbol{p})$ such that $(\boldsymbol{m}(\boldsymbol{p}), \boldsymbol{p})$ is a steady state. It remains to show that there is some $\boldsymbol{p} \in \boldsymbol{P}$ such that the tuple $(\boldsymbol{m}(\boldsymbol{p}), \boldsymbol{p})$ is also a partial equilibrium.

For every $\boldsymbol{p} \in \boldsymbol{P}$, we have $\beta_{HH}(\boldsymbol{m}(\boldsymbol{p}), \boldsymbol{p}) \leq \beta_{HL}(\boldsymbol{m}(\boldsymbol{p}), \boldsymbol{p})$. By Lemma 2.2, $V_{HH}(\boldsymbol{m}(\boldsymbol{p}), \boldsymbol{p}) > V_{HL}(\boldsymbol{m}(\boldsymbol{p}), \boldsymbol{p})$ because $\frac{u_{HH}}{u_{HL}} > 1$. Hence, H-agents behave optimally for every $\boldsymbol{p} \in \boldsymbol{P}$.

If $\boldsymbol{p} \in \boldsymbol{P}$ such that $p_{LLH} = 1$ and $p_{LHL} = 0$ is not an equilibrium strategy profile, then $V_{LL}(\boldsymbol{m}(\boldsymbol{p}), \boldsymbol{p}) > V_{LH}(\boldsymbol{m}(\boldsymbol{p}), \boldsymbol{p})$. Moreover, if $\boldsymbol{p} \in \boldsymbol{P}$ such that $p_{LLH} = 0$ and $p_{LHL} = 1$ is not an equilibrium strategy profile, then $V_{LH}(\boldsymbol{m}(\boldsymbol{p}), \boldsymbol{p}) > V_{LL}(\boldsymbol{m}(\boldsymbol{p}), \boldsymbol{p})$.

Note that $\boldsymbol{P}$ is connected (it is an L-shaped path parametrized by p_{LLH} and p_{LHL}). Moreover, V_{LL} and V_{LH} are both continuous functions of $(\boldsymbol{m}, \boldsymbol{p})$, and by Lemma 2.19, $\boldsymbol{m}(\boldsymbol{p})$ is continuous on $\boldsymbol{P}$. Thus, by the intermediate value theorem, there exists some $\boldsymbol{p}^* \in \boldsymbol{P}$ such that $V_{LL}(\boldsymbol{m}(\boldsymbol{p}^*), \boldsymbol{p}^*) = V_{LH}(\boldsymbol{m}(\boldsymbol{p}^*), \boldsymbol{p}^*)$ and $(\boldsymbol{m}(\boldsymbol{p}^*), \boldsymbol{p}^*)$ is a partial equilibrium. □

Remark 5. There are many ways to show that there is a continuum of steady state equilibria. For example, instead of varying p_{HLL}, we could have also varied p_{LHH} or p_{LLL}. The reason that we do not consider multiple variations at once, e.g., even varying p_{HHH} or relaxing that $p_{LLH} p_{LHL} = 0$, is that we can no longer guarantee a unique steady state. Of course, there is nothing suggesting that jointly considering these variations would not yield an even greater number of steady state equilibria.

Appendix 2.B Proofs: Augmented Model

2.B.1 Section 2.7

Proof of Theorem 2.11. Let $(\boldsymbol{m}, \boldsymbol{p})$ be a partial equilibrium. The proof narrows down the possible equilibrium behavior in several steps.

We write $\bar{i}$ for the type that is not type $i \in \{L, H\}$, i.e., $\{\bar{i}\} = \{L, H\} \setminus \{i\}$. Moreover, it will be convenient to let $r = 2(\delta + \lambda m) < \infty$ denote the maximal rate at which a match can possibly dissolve, where m is the total mass of agents in the market.

Step 1. For any ij-match, there exists a match duration $\bar{t} \geq 0$ after which both agents stay in the match, i.e., for all $t > \bar{t}$, $p_{ijL}(t) = 0$, $p_{ijH}(t) = 0$, $p_{jiL}(t) = 0$, and $p_{jiH}(t) = 0$.

For all $t \geq 0$,

$$V_{ij}(t) > \int_t^\infty e^{-r(s-t)} u_{ij}(s)\, ds \qquad \text{and} \qquad V_{ji}(t) > \int_t^\infty e^{-r(s-t)} u_{ji}(s)\, ds.$$

By our assumptions on productivity growth, both right-hand sides diverge to infinity as $t \to \infty$. Hence there is some $\bar{t} \geq 0$ such that for all $t > \bar{t}$,

$$V_{ij}(t) > \max\{V_{iL}(0), V_{iH}(0)\} \qquad \text{and} \qquad V_{ji}(t) > \max\{V_{jL}(0), V_{jH}(0)\}$$

since the right-hand sides are finite by assumption. The claim then follows from the definition of partial equilibrium.

Step 2. For any ij-match and any match duration $t^* > 0$, if
- $p_{ji\bar{i}}(t) = 0$ for all $t > t^*$ or $p_{ji\bar{i}}(t)$ is a cutoff strategy, and
- $p_{jii}(t) = 0$ for all $t > t^*$,

then there exists $\varepsilon > 0$ such that
(1) $V_{ij}(t) \gtrless V_{ij}(t^*)$ for all $t \gtrless t^*$, and $V_{ij}(t)$ is continuous at t^*,
(2) $p_{ijj}(t) = 0$ for all $t > t^* - \varepsilon$, and
(3) $V_{ij}(t^*) > V_{i\bar{j}}(0)$ implies $V_{ij}(t) > V_{i\bar{j}}(0)$ and $p_{ij\bar{j}}(t) = 0$ for all $t > t^* - \varepsilon$.

The premises imply $V_{ij}(t) \gtrless V_{ij}(t^*)$ for all $t \gtrless t^*$ in either of the two cases because flow utility $u_{ij}(t)$ is strictly increasing in match duration t and the rate $\beta_{ij}(t)$ at which an ij-match is dissolved by the j-agent satisfies $\beta_{ij}(t) \lesseqgtr \beta_{ij}(t^*)$ for all $t \gtrless t^*$.

For the continuity claim, note that

$$V_{ij}(t^*) > V_{ij}(t^* - \varepsilon) > \int_{t^*-\varepsilon}^{t^*} e^{-\bar{r}(s-t^*)} u_{ij}(s)\, ds + e^{-\bar{r}\varepsilon} V_{ij}(t^*).$$

The right-most expression converges to $V_{ij}(t^*)$ as $\varepsilon \downarrow 0$, thus $V_{ij}(t^* - \varepsilon) \to V_{ij}(t^*)$ as $\varepsilon \downarrow 0$. A similar argument shows that $V_{ij}(t^* + \varepsilon) \to V_{ij}(t^*)$ as $\varepsilon \downarrow 0$, hence $V_{ij}(t)$ is continuous at t^*.

Implications (2) and (3) follows immediately from the continuity of $V_{ij}(t)$ at t^* and the definition of partial equilibrium.

Step 3. For any ij-match, there exists a match duration $t^* \geq 0$ after which both agents stay in the match such that $p_{ij\bar{j}}$ or $p_{ji\bar{i}}$ is a cutoff strategy with cutoff t^* (or both).

By combining the first two steps, in particular implications (2) and (3) in Step 2, we can find some $t^* \geq 0$ after which both agents stay in the match such that $t^* = 0$ or $V_{ij}(t^*) \leq \max\{V_{iL}(0), V_{iH}(0)\}$ or $V_{ji}(t^*) \leq \max\{V_{jL}(0), V_{jH}(0)\}$. If $t^* = 0$, then we trivially establish Step 3. Thus, without loss of generality, suppose $V_{ij}(t^*) \leq \max\{V_{iL}(0), V_{iH}(0)\}$.

Since the j-agent stays in the match after t^*, the first implication in Step 2 yields $V_{ij}(t^*) > V_{ij}(t)$ for all $t < t^*$, thus $V_{ij}(t^*) \leq V_{i\bar{j}}(0)$. Again, by the first implication in Step 2, $V_{ij}(t) < V_{i\bar{j}}(0)$ for all $t < t^*$. Thus, the definition of partial equilibrium implies that $p_{ij\bar{j}}$ is a cutoff strategy.

Step 4. For all $t > 0$, $p_{HHH}(t) = 0$, and $p_{HHL}(t) = 0$.

By the previous step, $p_{HHL}(t)$ is a cutoff strategy with cutoff t^*, and both agents in an HH-match stay after t^*. Since $u_{HH}(t) > u_{HL}(t)$ for all $t \geq 0$, flow utility is increasing, and both agents stay in the match after t^*, we have $V_{HH}(t^*) > V_{HL}(0)$. Thus, by the definition of partial equilibrium, $t^* = 0$. Moreover, since both agents stay in the match after t^*, $p_{HHH}(t) = 0$ for all $t > 0$.

Step 5. For all $t > 0$, $p_{LLL}(t) = 0$, and $p_{LLH}(t)$ is a cutoff strategy with cutoff t_{LLH}.

By Step 3, $p_{LLH}(t)$ is a cutoff strategy with cutoff t^*, and both agents in an LL-match stay after t^*. For the sake of contradiction, suppose $p_{LLL}(t) > 0$ for some $t \in (0, t^*]$. Let $t' = \sup\{t \in (0, t^*] \mid p_{LLL}(t) > 0\}$. The second implication in Step 2 immediately contradicts the definition of t'.

Step 6. The strategies $p_{LHL}(t)$ and $p_{HLH}(t)$ are cutoff strategies with cutoffs t_{LHL} and t_{HLH}, respectively, where $t_{LHL} < t_{HLH}$. Moreover, for all $t > 0$, $p_{LHH}(t) = 0$ and $p_{HLL}(t) = 0$.

By Step 3, there exists a match duration t^* such that both agents in an LH-match stay in the match after t^* and such that $p_{HLH}(t)$ or $p_{LHL}(t)$ is a cutoff strategy with cutoff t^*. However, $p_{LHL}(t)$ cannot be a cutoff strategy with cutoff t^*: we have $V_{LH}(t^*) > V_{LL}(0)$ because both agents in an LH-match stay after t^* and an LH-match yields a higher flow utility for L-agents than an LL-match. Thus, $p_{HLH}(t)$ is a cutoff strategy with cutoff $t_{HLH} = t^*$.

In analogy to Step 3, by using the second and third implication in Step 2, we can find a match duration $t_{LHL} < t_{HLH}$ for an LH-match such that the L-agent stays after t_{LHL}, $p_{HLL}(t) = 0$ for all $t > t_{LHL}$, and $V_{LH}(t_{LHL}) \leq \{V_{LL}(0), V_{LH}(0)\}$ unless $t_{LHL} = 0$. If $t_{LHL} = 0$, we are done, thus suppose $t_{LHL} > 0$.

The first implication of Step 2 yields $V_{LH}(t_{LHL}) > V_{LH}(t)$ for all $t < t_{LHL}$, thus $V_{LH}(t_{LHL}) \leq V_{LL}(0)$. Again, by the first implication in Step 2, $V_{LH}(t) < V_{LL}(0)$ for all $t < t_{LHL}$. Thus, the definition of partial equilibrium implies that $p_{LHL}(t)$ is a cutoff strategy with cutoff $t_{LHL} < t_{HLH}$. Finally, by the same argument as in Step 5, we have $p_{LHH}(t) = 0$ and $p_{HLL}(t) = 0$ for all $t > 0$.

Step 7. $V_{ij}(t)$ is strictly increasing and continuous in match duration t for all $i, j \in \{L, H\}$.

Follows immediately from Steps 4-6 and the first implication in Step 2.

Step 8. For each cutoff t_{ijk}, if $t_{ijk} > 0$, then t_{ijk} is the unique solution to $V_{ij}(t_{ijk}) = V_{ik}(0)$.

Follows immediately from Step 7 and the definition of partial equilibrium.

Step 9. $t_{LLH} = 0$ or $t_{LHL} = 0$.

Suppose $t_{LLH} > 0$. By Steps 7 and 8, $V_{LL}(0) < V_{LL}(t_{LLH}) = V_{LH}(0) < V_{LH}(t_{LHL})$. Again by Steps 7 and 8, $t_{LHL} = 0$. □

2.B.2 Section 2.8

Balance conditions. Let $(\boldsymbol{m}, \boldsymbol{p})$ be a partial equilibrium. Then, using the partial equilibrium characterization in Theorem 2.11, the balance conditions for $(\boldsymbol{m}, \boldsymbol{p})$ to be a steady state are as follows; as before, the inflows will be on the left-hand side and the outflows on the right-hand side. For singles, we have

$$\begin{aligned} \eta_L + \delta(A+B+C+D+E) + \lambda D m_{H\emptyset} + \lambda(A+B)(m_{H\emptyset}+A+B) \\ = \delta m_{L\emptyset} + \lambda m_{L\emptyset}(m_{L\emptyset}+m_{H\emptyset}+A) \end{aligned} \tag{2.B.1}$$

$$\begin{aligned} \eta_H + \delta(A+B+C+m_{HH}) + \lambda A(m_{L\emptyset}+A) \\ = \delta m_{H\emptyset} + \lambda m_{H\emptyset}(m_{L\emptyset}+m_{H\emptyset}+A+B+D). \end{aligned} \tag{2.B.2}$$

For HH-matches, we have

$$\lambda(m_{H\emptyset}+A+B)^2 = 2\delta m_{HH}. \tag{2.B.3}$$

For LH-matches, i.e., A, B, and C, respectively, we have

$$\begin{aligned} &\lambda m_{H\emptyset}(m_{L\emptyset}+D) \\ &\quad = 2\delta A + \lambda A(m_{L\emptyset}+A) + \lambda A(m_{H\emptyset}+A+B) \\ &\qquad + \lambda m_{H\emptyset}(m_{L\emptyset}+D) q_{LH}(\boldsymbol{m},\boldsymbol{p};t_{LHL}) \end{aligned} \tag{2.B.4}$$

$$\begin{aligned} &\lambda m_{H\emptyset}(m_{L\emptyset}+D) q_{LH}(\boldsymbol{m},\boldsymbol{p};t_{LHL}) \\ &\quad = 2\delta B + \lambda B(m_{H\emptyset}+A+B) + \lambda m_{H\emptyset}(m_{L\emptyset}+D) q_{LH}(\boldsymbol{m},\boldsymbol{p};t_{HLH}) \end{aligned} \tag{2.B.5}$$

$$\lambda m_{H\emptyset}(m_{L\emptyset}+D) q_{LH}(\boldsymbol{m},\boldsymbol{p};t_{HLH}) = 2\delta C \tag{2.B.6}$$

where the survival function of an LH-match reads

$$q_{LH}(\boldsymbol{m},\boldsymbol{p};t)=\begin{cases}\exp\left(-(2\delta+\lambda(m_{L\emptyset}+A)+\lambda(m_{H\emptyset}+A+B))t\right) & 0\le t\le t_{LHL}\\ \exp\left(-\lambda(m_{L\emptyset}+A)t_{LHL}-(2\delta+\lambda(m_{H\emptyset}+A+B))t\right) & t_{LHL}\le t\le t_{HLH}\\ \exp\left(-\lambda(m_{L\emptyset}+A)t_{LHL}-\lambda(m_{H\emptyset}+A+B))t_{HLH}-2\delta t\right) & t\ge t_{HLH}.\end{cases}$$

For LL-matches, i.e., D and E, respectively, we have

$$\lambda(m_{L\emptyset}+A)^2 \;=\; 2\delta D+2\lambda Dm_{H\emptyset}+\lambda(m_{L\emptyset}+A)^2q_{LL}(\boldsymbol{m},\boldsymbol{p};t_{LLH}) \tag{2.B.7}$$

$$\lambda(m_{L\emptyset}+A)^2q_{LL}(\boldsymbol{m},\boldsymbol{p};t_{LLH}) \;=\; 2\delta E \tag{2.B.8}$$

where the survival function of an LL-match reads

$$q_{LL}(\boldsymbol{m},\boldsymbol{p};t)=\begin{cases}\exp\left(-(2\delta+2\lambda m_{H\emptyset})t\right) & 0\le t\le t_{LLH}\\ \exp\left(-2\lambda m_{H\emptyset}t_{LLH}-2\delta t\right) & t\ge t_{LLH}.\end{cases}$$

We now prove some preliminary results.

Lemma 2.23. *The expected continuation utilities $V_{i\emptyset}(\boldsymbol{m},\boldsymbol{p})$, $V_{iL}(\boldsymbol{m},\boldsymbol{p};t)$, and $V_{iH}(\boldsymbol{m},\boldsymbol{p};t)$ are continuous in $(\boldsymbol{m},\boldsymbol{p})$ for all $i\in\{L,H\}$ and $t\ge 0$.*[21]

Proof. As in the proof of Lemma 2.2 for the baseline model, the values $V_{i\emptyset}(\boldsymbol{m},\boldsymbol{p})$, $V_{iL}(\boldsymbol{m},\boldsymbol{p};0)$, and $V_{iH}(\boldsymbol{m},\boldsymbol{p};0)$ are the unique solution to a system of linear equations, namely (2.4) and (2.13). The coefficients of this system are continuous in $(\boldsymbol{m},\boldsymbol{p})$; hence, the unique solution is continuous in $(\boldsymbol{m},\boldsymbol{p})$ by the same argument as in the proof of Lemma 2.18. Given the continuity of $V_{iL}(\boldsymbol{m},\boldsymbol{p};0)$ and $V_{iH}(\boldsymbol{m},\boldsymbol{p};0)$, (2.13) immediately implies the continuity of $V_{iL}(\boldsymbol{m},\boldsymbol{p};t)$ and $V_{iH}(\boldsymbol{m},\boldsymbol{p};t)$. □

Lemma 2.24. *For any pair of single masses $(m_{L\emptyset},m_{H\emptyset})$, the following are true:*
(1) Given cutoffs $\boldsymbol{t}=(t_{LLH},t_{LHL},t_{HLH})$ as in the statement of Theorem 2.11, there is a unique solution (A,B,C,D,E,m_{HH}) to the balance conditions (2.B.3)-(2.B.8), i.e., all balance conditions except those of singles. For this solution, $AD=0$. Moreover, the solution is continuous in $\boldsymbol{t}$.
(2) Given masses (A,B,C,D,E) such that $AD=0$ and such that the aggregate balance conditions for m_{LH} and m_{LL}, i.e., the sums (2.B.4)+(2.B.5)+(2.B.6) and

21. As a technical aside, we must specify a norm for the space of all strategy profiles $\boldsymbol{p}$ when referring to the continuity of functions on that space. By Theorem 2.11, one may identify strategies $\boldsymbol{p}$ with cutoffs $\boldsymbol{t}$ and simply consider the Euclidean norm. However, we will sometimes consider infinite cutoffs, i.e., cutoffs in the extended reals. For this reason, we take the more general norm

$$||\boldsymbol{p}||=\max_{i,j,k\in\{L,H\}}\int_0^\infty p_{ijk}(t)e^{-2\delta t}.$$

(2.B.7)+(2.B.8), respectively, are satisfied, there are unique cutoffs $\mathbf{t}$ *as in the statement of Theorem 2.11 such the balance conditions (2.B.4)-(2.B.8) for* (A,B,C,D,E) *are satisfied. Moreover, the cutoffs* $\mathbf{t}$ *are continuous in* (A,B,C,D,E).

Proof. For both statements, the continuity claim follows immediately from the continuity of the balance conditions and the fact that there is a unique solution in either case (cf. the proof of Lemma 2.18).

For the first statement, note that m_{HH} is uniquely determined from (A,B,C,D,E) by (2.B.3). Moreover, by (2.B.6) and (2.B.8), C and E are uniquely determined from (A,B,D). We distinguish two cases: $t_{LHL}=0$ and $t_{LLH}=0$.

If $t_{LHL}=0$, then $A=0$ by (2.B.4). Consequently, D is uniquely determined from (2.B.7). Moreover, B is uniquely determined from (2.B.5). To see why, note that (2.B.5) describes the intersection of an increasing concave function in B (of the form $z_1(1-\exp-(z_2+B)t_{HLH})$ for some $z_1,z_2>0$) and a quadratic, hence convex function in B (of the form $B(z_3+z_4B)$ for some $z_3,z_4>0$). At $B=0$, the concave function is strictly positive, while the convex one is zero. Thus, there are exactly two candidate solutions, one negative and one positive.

If $t_{LLH}=0$, then $D=0$ by (2.B.7). Next, we show that A and B are uniquely determined by (2.B.4) and (2.B.5). Suppose not, i.e., there are two different pairs (A_1,B_1) and (A_2,B_2) solving these two equations. Without loss of generality, assume that $A_1\le A_2$. Note that (2.B.4) can be rewritten as

$$A \;=\; \lambda m_{H\emptyset}(m_{L\emptyset}+D)\int_0^{t_{LHL}} \exp\Big(-(2\delta+\lambda(m_{L\emptyset}+A)+\lambda(m_{H\emptyset}+A+B))t\Big)\;dt,$$

whose right-hand side is strictly decreasing in $2A+B$. Hence, we can conclude from $A_1\le A_2$ that $2A_1+B_1\ge 2A_2+B_2$, implying that $A_1+B_1\ge A_2+B_2$ and $B_1\ge B_2$. Similarly, (2.B.5) can be expressed as

$$\begin{aligned} &B\exp\Big((2\delta+\lambda(m_{L\emptyset}+A)+\lambda(m_{H\emptyset}+A+B))t_{LHL}\Big)\\ &\qquad=\;\lambda m_{H\emptyset}(m_{L\emptyset}+D)\int_{t_{LHL}}^{t_{HLH}} \exp\Big(-(2\delta+\lambda(m_{H\emptyset}+A+B))(t-t_{LHL})\Big)\;dt,\end{aligned}$$

whose right-hand side is strictly decreasing in $A+B$. Consequently, $B_1\ge B_2$ and $A_1+B_1\ge A_2+B_2$ yield $A_1+B_1\le A_2+B_2$. Therefore, we obtain that $A_1+B_1=A_2+B_2$, and thus $A_1=A_2$ and $B_1=B_2$.

The second statement follows immediately from the balance conditions (2.B.4)-(2.B.8) given that survival functions are strictly decreasing in match duration and the assumption that the aggregate balance conditions for m_{LH} and m_{LL} are satisfied. $\square$

Lemma 2.25. *Suppose* $(\mathbf{m},\mathbf{t})$ *and* $(\mathbf{m}',\mathbf{t}')$ *are two steady state equilibria.*
(1) If $t_{LHL}=t'_{LHL}=0$ *and* $B\le B'$*, then* $V_{Lj}(\mathbf{m},\mathbf{t};0)\ge V_{Lj}(\mathbf{m}',\mathbf{t}';0)$ *for all* $j\in\{L,H,\emptyset\}$.

(2) *If* $t_{LLH} = t'_{LLH} = 0$, $A \le A'$ *and* $A + B \ge A' + B'$, *then* $V_{Lj}(\boldsymbol{m}, \boldsymbol{t}; 0) \le V_{Lj}(\boldsymbol{m}', \boldsymbol{t}'; 0)$ *for all* $j \in \{L, H, \emptyset\}$.

Proof. To see the first statement, let $(B_t)_{t=0}^{\infty}$ be some sequence of masses, and let $\hat{V}_{L,j}((B_t)_t)$ be an L-agents continuation payoff who just entered a match with a j-agent, where her current partner (tth future partner) meets an H-agent in an LH-match who is willing to switch to an H-agent at rate λB_0 (λB_t), conditional on the L-agent and her current as well as all future partners behaving optimally given that constraint. In the steady state equilibrium $(\boldsymbol{m}, \boldsymbol{t})$ ($(\boldsymbol{m}', \boldsymbol{t}')$), all her partner meet such agents at rate λB ($\lambda B'$). Moreover, she is worse off if she is left at a faster rate ($\lambda(m_{H\emptyset} + B')$ instead of $\lambda(m_{H\emptyset} + B)$) in a match she does not want to leave (i.e., in a LH-match). By construction, we obtain that

$$\begin{aligned} V_{Lj}(\boldsymbol{m}, \boldsymbol{t}; 0) \;=\; \hat{V}_{L,j}((B)_t) \;&\ge\; \hat{V}_{L,j}((B', B, B, \ldots)) \\ &\ge\; \hat{V}_{L,j}((B', B', B, B, \ldots)) \\ &\ge\; \ldots \;\ge\; \hat{V}_{L,j}((B')_t) \;=\; V_{Lj}(\boldsymbol{m}', \boldsymbol{t}'; 0). \end{aligned}$$

For the second statement, let $(A_t, B_t)_{t=0}^{\infty}$ be some sequence of masses, and let $\hat{V}_{L,j}((A_t, B_t)_t)$ be an L-agents continuation payoff who just entered a match with a j-agent, where she meets some L-agent in an LH-match willing to switch to an L-agent at rate λA_0 (λA_t) while matched with her current partner (tth future partner), and her current partner (tth future partner) meets an H-agent in an LH-match willing to switch to an H-agent at rate $\lambda(A_0 + B_0)$ ($\lambda(A_t + B_t)$), conditional on the L-agent and her current as well as all future partners behaving optimally given these constraints. Now she is better off if she meets agents willing to rematch with her at a faster rate, and is left at a slower rate in any match. Consequently, we get

$$\begin{aligned} V_{Lj}(\boldsymbol{m}, \boldsymbol{t}; 0) \;=\; \hat{V}_{L,j}((A, B)_t) \;&\le\; \hat{V}_{L,j}(((A', B'), (A, B), (A, B), \ldots)) \\ &\le\; \hat{V}_{L,j}(((A', B'), (A', B'), (A, B), (A, B) \ldots)) \\ &\le\; \ldots \;\le\; \hat{V}_{L,j}((A', B')_t) \;=\; V_{Lj}(\boldsymbol{m}', \boldsymbol{t}'; 0). \end{aligned}$$

□

Proof of Theorem 2.12 (Existence). Our proof strategy is similar to that of other existence results in the literature: we construct a continuous mapping that reflects best-response and steady state dynamics and argue that a fixed point of this mapping corresponds to a steady state equilibrium. Specifically, consider a mapping

$$T : \boldsymbol{q} \overset{(1)}{\mapsto} (\boldsymbol{m}, \boldsymbol{t}) \overset{(2)}{\mapsto} (\hat{\boldsymbol{m}}, \hat{\boldsymbol{t}}) \overset{(3)}{\mapsto} \hat{\boldsymbol{q}},$$

where $\boldsymbol{q} = (q_{LLH}, q_{LHL}, q_{HLH})$ and $q_{ijk} \in [0, 1]$ is the share of i-agents matched with j-agents that accept to match with k-agents among all i-agents matched with j-agents. As before, $\boldsymbol{m}$ is a tuple of masses and $\boldsymbol{t} = (t_{LLH}, t_{LHL}, t_{HLH})$ is a profile of cutoffs. For the domain of inputs $\boldsymbol{q}$, we require $q_{LLH}q_{LHL} = 0$ and $q_{LHL} \leq q_{HLH}$. Note for later that this domain is homeomorphic to a two-dimensional compact convex set (it is a folded trapezoid). For each of the steps (1)-(3), we will now specify a well-defined and continuous mapping.

First, given rematching shares $\boldsymbol{q}$, we construct a steady state $(\boldsymbol{m}, \boldsymbol{t})$ that is continuous in $\boldsymbol{q}$. Note that in the baseline model, we may identify rematching shares $\boldsymbol{q}$ with strategies $\boldsymbol{p}$. Thus, by Lemma 2.19, there exist unique steady state masses $\boldsymbol{m}'$ consistent with $\boldsymbol{q}$. Moreover, $\boldsymbol{m}'$ is continuous in $\boldsymbol{q}$. Now, let $\boldsymbol{m} = (\boldsymbol{m}', A, B, C, D, E)$, where

$$
\begin{aligned}
A &= q_{LHL}m_{LH} \\
B &= (q_{HLH} - q_{LHL})m_{LH}, \\
C &= (1 - q_{HLH})m_{LH} \\
D &= q_{LLH}m_{LL} \\
E &= (1 - q_{LLH})m_{LL}.
\end{aligned}
$$

Given the above definitions, observe that the aggregate balance condition for m_{LH}, i.e., (2.B.4)+(2.B.5)+(2.B.6), coincides with the balance condition (2.A.5) for m_{LH} in the baseline model. Similarly, the aggregate balance condition for m_{LL}, i.e., (2.B.7)+(2.B.8), coincides with the balance condition (2.A.4) for m_{LL} in the baseline model. Thus, by Lemma 2.24, there are unique cutoffs $\boldsymbol{t}$ such that $(\boldsymbol{m}, \boldsymbol{t})$ is a steady state. Moreover, $\boldsymbol{t}$ is continuous in $\boldsymbol{m}$, hence in $\boldsymbol{q}$.

Second, given the steady state $(\boldsymbol{m}, \boldsymbol{t})$, we determine "best-response" cutoffs $\hat{\boldsymbol{t}}$ and implied new masses $\hat{\boldsymbol{m}}$ (that are not necessarily part of a steady state). Let $\hat{t}_{HLH}$ be the unique solution to (2.14). By the 7th step in the proof of Theorem 2.11, there are unique cutoffs $\hat{t}_{LLH}$ and $\hat{t}_{LHL}$ that solve

$$
\begin{aligned}
V_{LL}(\boldsymbol{m}, (t_{LLH}, t_{LHL}, \hat{t}_{HLH})); \hat{t}_{LLH}) &= V_{LH}(\boldsymbol{m}, (t_{LLH}, t_{LHL}, \hat{t}_{HLH})); 0) \\
V_{LL}(\boldsymbol{m}, (t_{LLH}, t_{LHL}, \hat{t}_{HLH})); 0) &= V_{LH}(\boldsymbol{m}, (t_{LLH}, t_{LHL}, \hat{t}_{HLH})); \hat{t}_{LHL}).
\end{aligned}
$$

Clearly, only one equation has a solution and for the other, we set the respective cutoff to zero; thus, $\hat{t}_{LLH}\hat{t}_{LHL} = 0$. Moreover, $\hat{t}_{LHL} < \hat{t}_{HLH}$ by (2.13) and the assumption that $u_{LL}(t) < u_{LH}(t)$ for all $t \geq 0$. It follows from Lemma 2.23 that $\hat{\boldsymbol{t}} = (\hat{t}_{LLH}, \hat{t}_{LHL}, \hat{t}_{HLH})$ is continuous in $(\boldsymbol{m}, \boldsymbol{t})$. By Lemma 2.24, given $\hat{t}$ and $\boldsymbol{m}'$, we can uniquely determine masses $(\hat{A}, \hat{B}, \hat{C}, \hat{D}, \hat{E})$, and these masses are continuous in $\hat{\boldsymbol{t}}$ and $\boldsymbol{m}$, hence in $\boldsymbol{t}$ and $\boldsymbol{m}$. Let $\hat{\boldsymbol{m}} = (\boldsymbol{m}', \hat{A}, \hat{B}, \hat{C}, \hat{D}, \hat{E})$.

Third, let $\hat{\boldsymbol{q}} = (\hat{q}_{LLH}, \hat{q}_{LHL}, \hat{q}_{HLH})$ be such that $\hat{q}_{LLH} = \hat{D}/m_{LL}$, $\hat{q}_{LHL} = \hat{A}/m_{LH}$, and $\hat{q}_{HLH} = (\hat{A} + \hat{B})/m_{LH}$. Since $\hat{t}_{LLH}\hat{t}_{LHL} = 0$, we have $\hat{q}_{LLH}\hat{q}_{LHL} = 0$. Moreover, $\hat{q}_{LLH} < \hat{q}_{HLH}$. Thus, $\hat{\boldsymbol{q}}$ is again in the domain of inputs to the composite mapping T.

By Brouwer's fixed point theorem, the continuous mapping T has a fixed point. Lemma 2.24 implies that at a fixed point, we have $(\boldsymbol{m}, \boldsymbol{t}) = (\hat{\boldsymbol{m}}, \hat{\boldsymbol{t}})$. By construction, the steady state $(\boldsymbol{m}, \boldsymbol{t})$ is then also a partial equilibrium. □

Proof of Theorem 2.12 (Identification). Fix some $(m_{L\emptyset}, m_{H\emptyset})$. As shown in the existence part of the proof, there exist masses (A, B, C, D, E, m_{HH}) and cutoffs $\boldsymbol{t}$ satisfying the balance conditions (2.B.3)-(2.B.8) and the conditions for a partial equilibrium characterized in Theorem 2.11. Then, the balance conditions (2.B.1) and (2.B.2) yield inflows (η_L, η_H) so that $(\boldsymbol{m}, \boldsymbol{t})$ is a steady state equilibrium given these inflows.

To prove uniqueness, suppose there exist two different steady state equilibria $(\boldsymbol{m}, \boldsymbol{t})$ and $(\boldsymbol{m}', \boldsymbol{t}')$ (with corresponding inflows (η_L, η_H) and (η_L', η_H'), respectively) that sustain the same single masses $(m_{L\emptyset}, m_{H\emptyset})$. We distinguish three cases: $t_{LHL} = t_{LHL}' = 0$, $t_{LLH} = t_{LLH}' = 0$ and $t_{LHL} = t_{LLH}' = 0$.

If $t_{LHL} = t_{LHL}' = 0$, then $A = A' = 0$ by (2.B.4). Without loss of generality, suppose $t_{LLH} \leq t_{LLH}'$. If $t_{LLH} = t_{LLH}'$, then $\boldsymbol{m} = \boldsymbol{m}'$ by (2.B.3)-(2.B.8) so that $(\boldsymbol{m}, \boldsymbol{t})$ and $(\boldsymbol{m}', \boldsymbol{t}')$ coincide. Hence, $t_{LLH} < t_{LLH}'$. Then, $D < D'$ by (2.B.7). Moreover, $B < B'$ by (2.B.5). To see this, recall that (2.B.5) describes the intersection concave function $z_1(1 - \exp -(z_2 + B)t_{HLH})$ the convex function $B(z_3 + z_4 B)$, which is increasing on the positive real line. Increasing D implies an upward shift of the concave function: If $D < D'$, then $z_1 < z_1'$. Consequently, the positive intersection of the two functions increases; thus, $B < B'$. By the 7th step of the proof of Theorem 2.11, $t_{LLH} < t_{LLH}'$ implies $V_{LL}(\boldsymbol{m}, \boldsymbol{t}; t_{LLH}) < V_{LL}(\boldsymbol{m}, \boldsymbol{t}; t_{LLH}')$. Furthermore, $B < B'$ yields $V_{LH}(\boldsymbol{m}, \boldsymbol{t}; 0) \geq \hat{V}_{LH}((B', B, B, \ldots))$ and $V_{L\emptyset}(\boldsymbol{m}, \boldsymbol{t}; 0) \geq V_{L\emptyset}(\boldsymbol{m}', \boldsymbol{t}'; 0)$ by Lemma 2.25. Finally, observe that

$$
\begin{aligned}
&V_{LL}(\boldsymbol{m}, \boldsymbol{t}; t_{LLH}) - V_{LH}(\boldsymbol{m}, \boldsymbol{t}; 0) \\
&\quad < V_{LL}(\boldsymbol{m}, \boldsymbol{t}; t_{LLH}') - \hat{V}_{LH}((B', B, B, \ldots)) \\
&\quad = \int_{t_{LLH}'}^{\infty} \big(u_{LL}(t) + \delta V_{L\emptyset}(\boldsymbol{m}, \boldsymbol{t}; 0)\big) e^{-2\delta(t - t_{LLH}')}\, dt - \Bigg[\int_{t_{HLH}}^{\infty} \delta V_{L\emptyset}(\boldsymbol{m}, \boldsymbol{t}; 0) e^{-2\delta t}\, dt \\
&\qquad - \int_{0}^{t_{HLH}} \big(u_{LH}(t) + \big(\delta + \lambda(m_{H\emptyset} + B')\big) V_{L\emptyset}(\boldsymbol{m}, \boldsymbol{t}; 0)\big) e^{-(2\delta + \lambda(m_{H\emptyset} + B'))t}\, dt \Bigg] \\
&\quad \leq \int_{t_{LLH}'}^{\infty} \big(u_{LL}(t) + \delta V_{L\emptyset}(\boldsymbol{m}', \boldsymbol{t}'; 0)\big) e^{-2\delta(t - t_{LLH}')}\, dt - \Bigg[\int_{t_{HLH}}^{\infty} \delta V_{L\emptyset}(\boldsymbol{m}', \boldsymbol{t}'; 0) e^{-2\delta t}\, dt \\
&\qquad - \int_{0}^{t_{HLH}} \big(u_{LH}(t) + \big(\delta + \lambda(m_{H\emptyset} + B')\big) V_{L\emptyset}(\boldsymbol{m}', \boldsymbol{t}'; 0)\big) e^{-(2\delta + \lambda(m_{H\emptyset} + B'))t}\, dt \Bigg] \\
&\quad = V_{LL}(\boldsymbol{m}', \boldsymbol{t}'; t_{LLH}') - V_{LH}(\boldsymbol{m}', \boldsymbol{t}'; 0),
\end{aligned}
$$

where the weak inequality is due to $V_{L\emptyset}(\boldsymbol{m}, \boldsymbol{t}; 0) \geq V_{L\emptyset}(\boldsymbol{m}, \boldsymbol{t}; 0)$. This yields a contradiction because

$$V_{LL}(\boldsymbol{m},\boldsymbol{t};t_{LLH})-V_{LH}(\boldsymbol{m},\boldsymbol{t};0) \;=\; 0 \;=\; V_{LL}(\boldsymbol{m}',\boldsymbol{t}';t'_{LLH})-V_{LH}(\boldsymbol{m}',\boldsymbol{t}';0)$$

if both $(\boldsymbol{m},\boldsymbol{t})$ and $(\boldsymbol{m}',\boldsymbol{t}')$ are a steady state equilibrium.

If $t_{LLH}=t'_{LLH}=0$, then $D=D'=0$ by (2.B.7). Without loss of generality, suppose $t_{LHL}\leq t'_{LHL}$. If $t_{LHL}=t'_{LHL}$, then $\boldsymbol{m}=\boldsymbol{m}'$ by (2.B.3)-(2.B.8) so that $(\boldsymbol{m},\boldsymbol{t})$ and $(\boldsymbol{m}',\boldsymbol{t}')$ coincide. Hence, $t_{LHL}<t'_{LHL}$. By (2.B.4), we get $A<A'$ or $2A+B<2A'+B'$. Suppose $A\geq A'$, so $2A+B<2A'+B'$ implying $A+B<A'+B'$. Then, (2.B.5) yields $B>B'$—a contradiction as $A\geq A'$ and $A+B<A'+B'$ imply $B<B'$. Hence, $A<A'$. Now suppose $A+B<A'+B'$ and thus $2A+B<2A'+B'$. Summing up (2.B.4) and (2.B.5) yields

$$\begin{aligned}A+B =&\lambda m_{H\emptyset}(m_{L\emptyset}+D)\\ &\cdot\Bigg[\int_0^{t_{LHL}}\exp\Big(-(2\delta+\lambda(m_{L\emptyset}+A)+\lambda(m_{H\emptyset}+A+B))t\Big)\;dt\\ &+\int_{t_{LHL}}^{t_{HLH}}\exp\Big(-\lambda(m_{L\emptyset}+A)t_{LHL}-(2\delta+\lambda(m_{H\emptyset}+A+B))t\Big)\;dt\Bigg].\end{aligned} \tag{2.B.9}$$

The left-hand side is strictly higher with the parameters A, B and t_{LHL} than with A', B', so $A+B>A'+B'$—a contradiction. Finally, observe that

$$\begin{aligned}&V_{LH}(\boldsymbol{m},\boldsymbol{t};t_{LHL})-V_{LL}(\boldsymbol{m},\boldsymbol{t};0)\\ &<\;V_{LH}(\boldsymbol{m},\boldsymbol{t};t'_{LHL})-V_{LL}(\boldsymbol{m},\boldsymbol{t};0)\\ &=\Bigg[\int_{t'_{LHL}}^{t_{HLH}}\Big(u_{LH}(t)+(\delta+\lambda(m_{H\emptyset}+A+B))V_{L\emptyset}(\boldsymbol{m},\boldsymbol{t};0)\Big)e^{-(2\delta+\lambda(m_{H\emptyset}+A+B))(t-t'_{LHL})}\;dt\\ &\qquad+\int_{t_{HLH}}^{\infty}\Big(u_{LH}(t)+\delta V_{L\emptyset}(\boldsymbol{m},\boldsymbol{t};0)\Big)e^{-2\delta(t-t'_{LHL})}\;dt\Bigg]\\ &\quad-\int_0^{\infty}\Big(u_{LL}(t)+\delta V_{L\emptyset}(\boldsymbol{m},\boldsymbol{t};0)\Big)e^{-2\delta t}\;dt\\ &\leq\Bigg[\int_{t'_{LHL}}^{t_{HLH}}\Big(u_{LH}(t)+(\delta+\lambda(m_{H\emptyset}+A'+B'))V_{L\emptyset}(\boldsymbol{m}',\boldsymbol{t}';0)\Big)e^{-(2\delta+\lambda(m_{H\emptyset}+A'+B'))(t-t'_{LHL})}\;dt\\ &\qquad+\int_{t_{HLH}}^{\infty}\Big(u_{LH}(t)+\delta V_{L\emptyset}(\boldsymbol{m}',\boldsymbol{t}';0)\Big)e^{-2\delta(t-t'_{LHL})}\;dt\Bigg]\\ &\quad-\int_0^{\infty}\Big(u_{LL}(t)+\delta V_{L\emptyset}(\boldsymbol{m}',\boldsymbol{t}';0)\Big)e^{-2\delta t}\;dt\\ &=\;V_{LH}(\boldsymbol{m}',\boldsymbol{t}';t'_{LHL})-V_{LL}(\boldsymbol{m}',\boldsymbol{t}';0),\end{aligned}$$

where the strict inequality follows from strict monotonicity of V_{LL} (see 7th step of the proof of Theorem 2.11), and the weak inequality is due to $A+B>A'+B'$

and $V_{L\emptyset}(\boldsymbol{m},\boldsymbol{t};0) \le V_{L\emptyset}(\boldsymbol{m}',\boldsymbol{t}';0)$ (see Lemma 2.25). Again, we obtain that $(\boldsymbol{m},\boldsymbol{t})$ or $(\boldsymbol{m}',\boldsymbol{t}')$ cannot be a steady state equilibrium.

If $t_{LHL} = t'_{LLH} = 0$, then $t_{LLH}, t'_{LHL} > 0$ by the previous two cases and thus $0 = A < A'$ and $D > D' = 0$. From (2.B.9), we can conclude that $A + B > A' + B'$. Let $\boldsymbol{m}''$ be the tuple of masses satisfying the balance equations (2.B.3)-(2.B.8) (given the single masses $(m_{L\emptyset}, m_{H\emptyset})$) for the cutoffs $\boldsymbol{t}'' = (0, 0, t_{HLH})$, that is, $A'' = D'' = 0$, and $A + B > A'' + B'' > A' + B'$ by (2.B.9). Following similar arguments as in the proof of the first two cases, we obtain that

$$\begin{aligned} V_{LL}(\boldsymbol{m}',\boldsymbol{t}';0) - V_{LH}(\boldsymbol{m}',\boldsymbol{t}';t'_{LHL}) &< V_{LL}(\boldsymbol{m}'',\boldsymbol{t}'';0) - V_{LH}(\boldsymbol{m}'',\boldsymbol{t}'';0) \\ &< V_{LL}(\boldsymbol{m},\boldsymbol{t};t_{LLH}) - V_{LH}(\boldsymbol{m},\boldsymbol{t};0), \end{aligned}$$

contradicting the fact that both $(\boldsymbol{m},\boldsymbol{t})$ and $(\boldsymbol{m}',\boldsymbol{t}')$ form a steady state equilibrium. □

Appendix 2.C Proofs: Equilibrium Selection & Model Comparison

2.C.1 Section 2.9

Proof of Theorem 2.15. Let $(\boldsymbol{m},\boldsymbol{p})$ be a limit equilibrium; let $(\boldsymbol{u}^n)$ and $(\boldsymbol{m}^n,\boldsymbol{p}^n)$ be the associated sequences of flow utilities and steady state equilibria; let $(\boldsymbol{m}^*,\boldsymbol{p}^*)$ be a steady state equilibrium that corresponds to $(\boldsymbol{m},\boldsymbol{p})$. We will first show that $(\boldsymbol{m}^*,\boldsymbol{p}^*)$ is a coordination equilibrium.

By Theorem 2.11, $p^n_{HHL}(t) = 0$ and $p^n_{ijj}(t) = 0$ for all $i,j \in \{L,H\}$ and $t > 0$, hence $p_{HHL}(t) = 0$ and $p_{ijj}(t) = 0$ for all $i,j \in \{L,H\}$ and $t > 0$. Since $(\boldsymbol{m}^*,\boldsymbol{p}^*)$ corresponds to $(\boldsymbol{m},\boldsymbol{p})$, equation (2.15) implies $p^*_{HHL} = 0$, $p^*_{ijj} = 0$ for all $i,j \in \{L,H\}$, and

$$p^*_{LHL} = A/m_{LH}, \qquad p^*_{LLH} = D/m_{LH}, \qquad p^*_{HLH} = (A+B)/m_{LH}.$$

By Theorem 2.11, $A^n = 0$ or $D^n = 0$, hence $A = 0$ or $D = 0$ and, consequently, $p^*_{LHL} = 0$ or $p^*_{HLH} = 0$. Finally, equation (2.14) and the initial assumption that $u^n_{HL}(t)$ and $u^n_{HH}(t)$ converge to u^*_{HL} and u^*_{HH} for all $t \ge 0$, respectively, where $u^*_{HL} < u^*_{HH}$, imply that t^n_{HLH} diverges to infinity, hence $(A+B)^n$ converges to m_{LH}, thus $p^*_{HLH} = 1$. Altogether, we have shown that $(\boldsymbol{m}^*,\boldsymbol{p}^*)$ is a coordination equilibrium.

For the converse, let $(\boldsymbol{m}^*,\boldsymbol{p}^*)$ be a coordination equilibrium. We will construct a corresponding limit equilibrium.

Suppose $p^*_{LLH} = 0$. Fix an arbitrary sequence of strategies $(\boldsymbol{p}^n)$ for the baseline model such that $p^n_{LHL} < p^n_{HLH} < 1$ for all $n \in \mathbb{N}$, $p^n_{LHL} \to p^*_{LHL}$, $p^n_{HLH} \to p^*_{HLH} = 1$, and all other switching probabilities are as in $\boldsymbol{p}^*$. By Lemma 2.19, $\boldsymbol{m}^n = \boldsymbol{m}^n(\boldsymbol{p}^n)$ is the unique tuple of steady state masses associated with $\boldsymbol{p}^n$, and also by Lemma 2.19, $\boldsymbol{m}^n \to \boldsymbol{m}^*$. Let

$$A^n = p^n_{LHL} m^n_{LH},\ B^n = (p^n_{HLH} - p^n_{LHL}) m^n_{LH},\ C^n = m^n_{LH} - A^n - B^n,\ D^n = 0,\ E^n = m^n_{LL},$$

and with slight abuse of notation, we understand $\boldsymbol{m}^n$ to be augmented with the masses A^n through E^n when referring to the augmented model.

Now, according to Lemma 2.24, we can choose $\boldsymbol{t}^n = (t^n_{LHL}, t^n_{LLH}, t^n_{HLH})$ with $t^n_{LLH} = 0$ and $t^n_{LHL} < t^n_{HLH}$ such that the balance conditions for (A, B, C) are satisfied (equations (2.B.4) to (2.B.6)). Note that as $n \to \infty$, $C^n \to 0$, hence $t^n_{HLH} \to \infty$.

By construction, $(\boldsymbol{m}^n, \boldsymbol{t}^n)$ is a steady state of the augmented model. We will now construct flow utilities $\boldsymbol{u}^n = \left(u^n_{ij}(t)\right)_{i,j\in\{L,H\},\, t\geq 0}$ such that $(\boldsymbol{m}^n, \boldsymbol{t}^n)$ is also a partial equilibrium of the augmented model.

Let us first construct the flow utilities for L-agents. Let (a^n) be a sequence such that $0 < a^n < u^*_{LH} - u^*_{LL}$, $a^n \to 0$, and if flow utilities for L-agents are given by u^*_{LL} and $u^*_{LH} - a^n$ in the baseline model, then $V_{LL}(\boldsymbol{m}^n, \boldsymbol{p}^n) > V_{LH}(\boldsymbol{m}^n, \boldsymbol{p}^n)$. A sequence (a^n) with the desired properties exists because expected continuation are continuous in their arguments (see the proof of Lemma 2.22), $(\boldsymbol{m}^n, \boldsymbol{p}^n) \to (\boldsymbol{m}^*, \boldsymbol{p}^*)$, and $V_{LL}(\boldsymbol{m}^*, \boldsymbol{p}^*) \geq V_{LH}(\boldsymbol{m}^*, \boldsymbol{p}^*)$ since $p^*_{LLH} = 0$ by assumption. Also, let (b^n_1) be a any strictly positive sequence converging to 0. Using these sequences, define

$$
\begin{aligned}
u^n_{LL}(t) &= u^*_{LL} + b^n_1 t \\
u^n_{LH}(t) &= u^*_{LH} - a^n + b^n_2 t,
\end{aligned}
$$

where $b^n_2 > 0$ is chosen such that

$$
V_{LL}(\boldsymbol{m}^n, \boldsymbol{t}^n; 0) = V_{LH}(\boldsymbol{m}^n, \boldsymbol{t}^n; t^n_{LHL})
$$

if flow utilities for L-agents are given by $u^n_{LL}(t)$ and $u^n_{LH}(t)$ in the augmented model.

The existence (and uniqueness) of b^n_2 follows from an intermediate value argument. First, for large enough $b^n_2 > 0$, $V_{LL}(\boldsymbol{m}^n, \boldsymbol{p}^n; 0) < V_{LH}(\boldsymbol{m}^n, \boldsymbol{p}^n; t^n_{LHL})$. Second, for all small enough $b^n_2 > 0$, $V_{LL}(\boldsymbol{m}^n, \boldsymbol{t}^n; 0) > V_{LH}(\boldsymbol{m}^n, \boldsymbol{t}^n; 0)$ because $V_{LL}(\boldsymbol{m}^n, \boldsymbol{p}^n) > V_{LH}(\boldsymbol{m}^n, \boldsymbol{p}^n)$ in the baseline model and because under the strategies $\boldsymbol{p}^n$ for the baseline model, agents switch randomly, whereas under the constructed strategies $\boldsymbol{t}^n$ for the augmented model, agents switch deterministically—stay earlier and switch later—which directly lowers V_{LH} because agents discount future flow utility but only indirectly lowers V_{LL} through future flow utility from an LH-match. Thus, for small enough $b^n_2 > 0$, we also have $V_{LL}(\boldsymbol{m}^n, \boldsymbol{t}^n; 0) > V_{LH}(\boldsymbol{m}^n, \boldsymbol{t}^n; t^n_{LHL})$. Third, both expected continuation values V_{LL} and V_{LH} are continuous in b^n_2; hence the desired $b^n_2 > 0$ exists.

Let us now construct the flow utilities for H-agents:

$$
\begin{aligned}
u^n_{HL}(t) &= u^*_{HL} + \frac{u^*_{HH} - u^*_{HL}}{t^n_{HLH}} \min\{t, t^n_{HLH}\} + \frac{1}{n} \max\{t - t^n_{HLH}, 0\} \\
u^n_{HH}(t) &= u^*_{HH} + \frac{1}{n} t.
\end{aligned}
$$

With these flow utilities, we ensure by (2.14) that

$$V_{HH}(\boldsymbol{m}^n, \boldsymbol{t}^n; 0) = V_{HL}(\boldsymbol{m}^n, \boldsymbol{t}^n; t^n_{HLH}).$$

By construction, $(\boldsymbol{m}^n, \boldsymbol{t}^n)$ is then a steady state equilibrium given the flow utilities $\boldsymbol{u}^n$.

Finally, let us show that the flow utilities $\boldsymbol{u}^n$ converge pointwise to $\boldsymbol{u}^*$, which implies that $(\boldsymbol{m}, \boldsymbol{t}) = \lim_{n\to\infty}(\boldsymbol{m}^n, \boldsymbol{t}^n)$ is the desired limit equilibrium that corresponds to $(\boldsymbol{m}^*, \boldsymbol{p}^*)$ (where convergence of $\boldsymbol{t}^n$ is understood to be in the extended reals). It is clear that $u^n_{HL}(t) \to u^*_{HL}$ and $u^n_{HH}(t) \to u^*_{HH}$ for all $t \geq 0$ because $t^n_{HLH} \to \infty$.

For the flow utilities of L-agents, it remains to show that $b^n_2 \to 0$. If $t^n_{LHL} \to \infty$ and $\lim_{n\to\infty} b^n_2 > 0$, then $V_{LH}(\boldsymbol{m}^n, \boldsymbol{p}^n; t^n_{LHL})$ diverges to infinity, whereas $V_{LL}(\boldsymbol{m}^n, \boldsymbol{p}^n; 0)$ is finite, which contradicts the definition of b^n_2. If $\lim_{n\to\infty} t^n_{LHL} < \infty$, then the definition of t^n_{LHL} implies that $\lim_{n\to\infty} B^n > 0$, hence $p^*_{LHL} < 1$ and consequently, $V_{LH}(\boldsymbol{m}^*, \boldsymbol{p}^*) = V_{LL}(\boldsymbol{m}^*, \boldsymbol{p}^*)$ by the definition of partial equilibrium. However, if $\lim_{n\to\infty} b^n_2 > 0$ and flow utilities in the augmented model are given by $\lim_{n\to\infty} \boldsymbol{u}^n$, then $V_{LH}(\boldsymbol{m}^*, \boldsymbol{p}^*) = V_{LL}(\boldsymbol{m}^*, \boldsymbol{p}^*)$ implies $V_{LH}(\boldsymbol{m}, \boldsymbol{t}; t_{LHL}) > V_{LL}(\boldsymbol{m}, \boldsymbol{t}; 0)$ because $t^n_{HLH} \to \infty$, i.e., the behavior of H-agents in $(\boldsymbol{m}, \boldsymbol{t})$ is the same as in $(\boldsymbol{m}^*, \boldsymbol{p}^*)$, and the flow utility of an L-agent in an LH-match is strictly increasing in match duration whereas it is constant in an LL-match. The fact that $V_{LH}(\boldsymbol{m}, \boldsymbol{t}; t_{LHL}) > V_{LL}(\boldsymbol{m}, \boldsymbol{t}; 0)$ given flow utilities $\lim_{n\to\infty} \boldsymbol{u}^n$ implies that $V_{LH}(\boldsymbol{m}^n, \boldsymbol{t}^n; t^n_{LHL}) > V_{LL}(\boldsymbol{m}^n, \boldsymbol{t}^n; 0)$ for all sufficiently large $n \in \mathbb{N}$ by the continuity of expected continuation utilities in masses, strategies, match durations, and flow utilities; see Lemma 2.23 and the proof of Theorem 2.11. This contradicts the definition of b^n_2, thus $\lim_{n\to\infty} b^n_2 = 0$.

Recall that we have initially assumed $p^*_{LLH} = 0$; the constructions for $p^*_{LHL} = 0$ are similar. □

2.C.2 Section 2.10

Proof of Theorem 2.16. As a preliminary observation, note that the masses $\boldsymbol{m}$ for which there exists that a cutoff profile $\boldsymbol{t} = (t_{LLH}, t_{LHL}, t_{HLH})$ such that $(\boldsymbol{m}, \boldsymbol{t})$ is a steady state are all bounded away from zero due to search frictions; see the balance conditions (2.B.1)-(2.B.8). Let $\boldsymbol{M}$ denote the set of such masses.

For the first claim, suppose $t_{LHL} = 0$. Fix $t_{HLH} = \infty$. Consider any $\boldsymbol{m} \in \boldsymbol{M}$ with $A = 0$ and $D \geq 0$ and any cutoff $t_{LLH} \geq 0$. By (2.13), we then have

$$\begin{aligned}
V_{LL}(t_{LLH}) &= \int_0^\infty e^{-2\beta_{LL}s} u_{LL}(s + t_{LLH})\, ds + \frac{1}{2} V_{L\emptyset} \\
V_{LH}(0) &= \int_0^\infty e^{-(\beta_{LH}+\beta_{HL})s} u_{LH}(s)\, ds + \frac{\beta_{LH}}{\beta_{LH} + \beta_{HL}} V_{L\emptyset}
\end{aligned}$$

where $\beta_{LH} = \delta + \lambda(m_{H\emptyset} + m_{LH})$ and $\beta_{HL} = \delta$ are constant, and $\beta_{LL} = \delta$ is constant for match durations $s > t_{LLH}$.

Now, if $\int_0^\infty \big(u_{LH}(t) - u_{LL}(t)\big) e^{-2\delta t}\, dt$ is sufficiently small, then $V_{LL}(t_{LLH}) > V_{LH}(0)$ for all $\boldsymbol{m} \in \boldsymbol{M}$ with $A = 0$ and $D \geq 0$ and cutoffs $t_{LLH} \geq 0$. To see this, note that

$$\frac{\beta_{LH}}{\beta_{LH}+\beta_{HL}}V_{L\emptyset} \leq \left(1-\frac{\beta_{LH}}{\beta_{LH}+\beta_{HL}}\right)V_{L\emptyset}+\left(2\frac{\beta_{LH}}{\beta_{LH}+\beta_{HL}}-1\right)V_{LH}(0).$$

(If $V_{L\emptyset} > V_{LH}(0)$, then $V_{LL}(0) > V_{L\emptyset}$; thus, we would be done.) Using the above inequality, we can bound the difference $V_{LL}(t_{LLH}) - V_{LH}(0)$ from below by

$$\int_0^\infty e^{-2\beta_{LL}s}u_{LL}(s)\,ds - \frac{\beta_{LH}+\beta_{HL}}{2\beta_{HL}}\int_0^\infty e^{-(\beta_{LH}+\beta_{HL})s}u_{LH}(s)\,ds.$$

Multiplying the expression with $2\beta_{LL} = 2\beta_{HL}$ reveals that it is nothing but the difference of the expectations of two exponential random variables. Since β_{LH} is bounded away from $\beta_{LL} = \beta_{HL} = \delta$ for all $\boldsymbol{m} \in \boldsymbol{M}$ with $A = 0$ and $D \geq 0$, the difference is indeed strictly positive if $\int_0^\infty \big(u_{LH}(t) - u_{LL}(t)\big)e^{-2\delta t}\,dt$ is sufficiently small.

Recall that we have fixed $t_{HLH} = \infty$. However, by Lemma 2.23, i.e., the continuity of expected continuation utilities, we still have

$$V_{LL}(\boldsymbol{m},(t_{LLH},0,t_{HLH});t_{LLH}) > V_{LH}(\boldsymbol{m},(t_{LLH},0,t_{HLH});0)$$

for all $\boldsymbol{m} \in \boldsymbol{M}$ with $A = 0$ and $D \geq 0$, all $t_{LLH} \geq 0$, and all sufficiently large t_{HLH}, which, by (2.14), is tantamount to $\int_0^\infty \big(u_{HH}(t) - u_{HL}(t)\big)e^{-2\delta t}\,dt$ being sufficiently large. This establishes the first claim because by Theorem 2.11, $t_{LHL} = 0$ would not be optimal, implying that there cannot exist a non-assorting equilibrium, and by Theorem 2.12, there must then exist an assorting equilibrium.

For the second claim, suppose $t_{LLH} = 0$. Consider any $\boldsymbol{m} \in \boldsymbol{M}$ with $A \geq 0$ and $D = 0$ and any cutoffs $0 \leq t_{LHL} < t_{HLH}$. As in the proof of Theorem 2.11, let $r = 2(\delta + \lambda m) < \infty$ denote the maximal rate at which a match can possibly dissolve, where m is the total mass of agents in the market. By (2.13), we then have

$$\begin{aligned} V_{LL}(0) &= \int_0^\infty e^{-2\delta s}u_{LL}(s)\,ds + \frac{1}{2}V_{L\emptyset} \\ V_{LH}(t_{LHL}) &> \exp(-r(t_{HLH}-t_{LHL}))\int_0^\infty e^{-2\delta s}u_{LH}(s+t_{HLH})\,ds + \frac{1}{2}V_{L\emptyset}. \end{aligned}$$

Thus, we can bound the difference $V_{LL}(t_{LLH}) - V_{LH}(0)$ from below by

$$\exp(-r(t_{HLH}-t_{LHL}))\int_0^\infty e^{-2\delta s}u_{LH}(s+t_{HLH})\,ds - \int_0^\infty e^{-2\delta s}u_{LL}(s)\,ds.$$

By assumption, $u_{LH}(t) > u_{LL}(t)$ for all $t \geq 0$; hence, the difference is strictly positive for all sufficiently small $t_{HLH} > 0$, which, by (2.14), is tantamount to $\int_0^\infty \big(u_{HH}(t) - u_{HL}(t)\big)e^{-2\delta t}\,dt$ being sufficiently small. Alternatively, if $\int_0^\infty \big(u_{LH}(t) - u_{LL}(t)\big)e^{-2\delta t}\,dt$ is sufficiently large, then the above difference is also strictly positive for any $0 \leq t_{LHL} < t_{HLH}$, where t_{HLH} solves (2.14). This establishes the second claim because by Theorem 2.11, $t_{LLH} = 0$ would not be optimal, implying that there cannot exist an assorting equilibrium, and by Theorem 2.12, there must then exist a non-assorting equilibrium. □

References

Acharya, Sushant, and Shu Lin Wee. 2020. “On-the-job Search and the Productivity-Wage Gap.” CEPR Discussion Paper No. DP14430. [65, 68]

Albak, Karsten, and Bent E Sørensen. 1998. “Worker flows and job flows in Danish manufacturing, 1980-91.” *Economic Journal* 108 (451): 1750–71. [65]

Bartolucci, Cristian, and Ignacio Monzon. 2019. “Frictions lead to sorting: A partnership model with on-the-match search.” [69]

Becker, Gary S. 1964. *Human Capital.* University of Chicago Press. [66, 78]

Bobbio, Emmanuele. 2019. “Replacement Search and Wages.” Working Paper. [65, 68]

Burdett, Kenneth, Ryoichi Imai, and Randall Wright. 2004. “Unstable relationships.” *B.E. Journal of Macroeconomics* 1 (1): 1–42. [69]

Burgess, Simon, Julia Lane, and David Stevens. 2000. “Job flows, worker flows, and churning.” *Journal of Labor Economics* 18 (3): 473–502. [65]

Chade, Hector, Jan Eeckhout, and Lones Smith. 2017. “Sorting through search and matching models in economics.” *Journal of Economic Literature* 55 (2): 493–544. [66, 69]

Chen, Yiqun. 2021. “Team-specific human capital and team performance: Evidence from doctors.” *American Economic Review* 111 (12): 3923–62. [66]

Cornelius, Tracy J. 2003. “A search model of marriage and divorce.” *Review of Economic Dynamics* 6 (1): 135–55. [68, 69]

Diamond, Peter A, and Eric Maskin. 1979. “An equilibrium analysis of search and breach of contract, I: Steady states.” *Bell Journal of Economics,* 282–316. [70]

Dolado, Juan J., Marcel Jansen, and Juan F. Jimeno. 2009. “On-the-job search in a matching model with heterogeneous jobs and workers.” *Economic Journal* 119 (534): 200–228. [68]

Fallick, Bruce, and Charles A. Fleischman. 2004. “Employer-to-employer flows in the US labor market: The complete picture of gross worker flows.” [65]

Flinn, Christopher, Ahu Gemici, and Steven Laufer. 2017. “Search, matching and training.” *Review of Economic Dynamics* 25: 260–97. [87]

Galperin, Roman V, Oliver Hahl, Adina D Sterling, and Jerry Guo. 2020. “Too good to hire? Capability and inferences about commitment in labor markets.” *Administrative Science Quarterly* 65 (2): 275–313. [87]

Gautier, Pieter A., Coen N. Teulings, and Aico van Vuuren. 2010. “On-the-job search, mismatch and efficiency.” *Review of Economic Studies* 77 (1): 245–72. [68]

Goldmanis, Maris, and Korok Ray. 2019. “Until Death Do Us Part? The Marriage Model with Divorce.” [69]

Hall, Robert E. 1982. “The Importance of Lifetime Jobs in the US Economy.” *American Economic Review* 72 (4): 716–24. [67]

Jaravel, Xavier, Neviana Petkova, and Alex Bell. 2018. “Team-specific capital and innovation.” *American Economic Review* 108 (4-5): 1034–73. [66]

Kellogg, Ryan. 2011. “Learning by drilling: Interfirm learning and relationship persistence in the Texas oilpatch.” *Quarterly Journal of Economics* 126 (4): 1961–2004. [66]

Lazear, Edward P. 2009. “Firm-specific human capital: A skill-weights approach.” *Journal of Political Economy* 117 (5): 914–40. [66]

Lentz, Rasmus. 2010. “Sorting by search intensity.” *Journal of Economic Theory* 145 (4): 1436–52. [68]

Lentz, Rasmus, and Nicolas Roys. 2015. “Training and Search on the Job.” Working paper. [87]

Murphy, Kevin J, and Jan Zabojnik. 2007. "Managerial capital and the market for CEOs." Working paper. [65]

Parrino, Robert. 1997. "CEO turnover and outside succession a cross-sectional analysis." *Journal of Financial Economics* 46 (2): 165–97. [65]

Pissarides, Christopher A. 1994. "Search unemployment with on-the-job search." *Review of Economic Studies* 61 (3): 457–75. [69]

Rogerson, Richard, Robert Shimer, and Randall Wright. 2005. "Search-theoretic models of the labor market: A survey." *Journal of Economic Literature* 43 (4): 959–88. [68]

Schwartz, Jeremy. 2020. "Job competition, human capital, and the lock-in effect: can unemployment insurance efficiently allocate human capital." *B.E. Journal of Macroeconomics* 20 (1): 1–23. [69]

Shimer, Robert, and Lones Smith. 2000. "Assortative matching and search." *Econometrica* 68 (2): 343–69. [65, 66, 68, 70, 87, 89]

Smith, Lones. 2006. "The marriage model with search frictions." *Journal of Political Economy* 114 (6): 1124–44. [65, 66, 68–70, 87]

Stevenson, Betsey, and Justin Wolfers. 2007. "Marriage and divorce: Changes and their driving forces." *Journal of Economic Perspectives* 21 (2): 27–52. [65]

Topel, Robert. 1991. "Specific capital, mobility, and wages: Wages rise with job seniority." *Journal of Political Economy* 99 (1): 145–76. [66]

Chapter 3

Clustering with a Minimum Distance Constraint

3.1 Introduction

Cluster analysis is a method of grouping data records in order to identify similarity patterns within these data. The objects of a data set are divided into several groups, also called clusters, based on their observable characteristics. The idea is to select the groups to maximize intra-cluster homogeneity and inter-cluster heterogeneity with respect to the data characteristics. Clustering is a useful method to discern information about the structure and composition of data and is therefore used in many scientific fields like economics, physics, or medicine.

In this paper, I analyze a new constrained clustering problem: Given a data set, allocate its data points to clusters in order to minimize the sum of squared distances between the cluster centers and their assigned data points subject to a minimum distance between the cluster centers. Cluster assignment is probabilistic, meaning that a data point is not necessarily allocated to only one cluster, but can be assigned to several clusters with positive probability. The standard approach to clustering problems is applying an algorithm that yields a numerical solution, such as the classical k-means algorithm. In contrast, I use optimization techniques to characterize the optimal clustering analytically.

The trade-off between intra-cluster homogeneity and inter-cluster heterogeneity is a frequently discussed topic in the literature on cluster analysis, and it is closely related to the problem of determining the optimal number of clusters: Increasing the number of different clusters improves within-class similarity at the expense of making clusters less distinguishable. Without further investigation of the data record, it is a complex task to set this number. Imposing a lower bound on the distance between the cluster centroids ensures that the formed clusters are sufficiently heterogeneous. Indeed, the optimal number of clusters can be derived from the minimum distance parameter in my setting. Justification of an appropriate distance between clusters is more straightforward than choosing the number of clusters directly. For instance, one could select a reasonable minimum distance by evaluating the space covered by the data record: If the maximal distance between any pair of data points is rather small, suggesting that the data space is rather dense, it makes sense to choose a moderate, not too large minimum distance between the clusters. Hence, specifying a minimum distance can be used as a tool to find the best number of clusters.

Probabilistic clustering is a generalization of deterministic clustering, meaning that each data point is associated with exactly one cluster. An alternative interpretation of probabilistic clustering is that shares of data points are assigned to clusters. Probabilistic clustering is a promising strategy to handle data sets with small errors or uncertainties: Under such circumstances, the data characteristics might not be deterministic so that randomizing over different cluster assignments can become optimal.

The basic k-means clustering problem aims at minimizing the sum of squared distances without imposing any side constraints. Also, only deterministic assignments are considered. It is well known that the optimal clustering is a convex partition, that is, if two data points are assigned to the same cluster, then any data point on the line between these two points also belongs to that cluster. This result extends to the case with probabilistic clustering: If two data points are allocated to a cluster with positive probability, then any data point that lies in between belongs to that cluster with certainty. These probabilistic partitions are called monotone.

With the minimum distance constraint, this changes : An optimal partition belongs to the class of so-called bi-pooling partitions: If two data points are assigned to a cluster with positive probability, any data point in between is allocated to at most two clusters with positive probability, namely to this cluster or to one other cluster, which is the same for all such points. This result builds on optimization techniques in convex analysis (see Bertsekas, Nedic, and Ozdaglar (2003), for instance): A linear function defined on a compact, convex set attains its maximum at an extreme point[1] of that set. Building on the work of Chapter 1, I prove that any extreme point of the set of probabilistic partitions satisfying the minimum distance constraint is a bi-pooling partition. This finding suggests that monotone partitions are not necessarily globally optimal: While every monotone partition is a bi-pooling partition, the converse does not hold true. In fact, I show by example that there exist clustering problems that are solved by a non-monotone bi-pooling partition.

Optimality of bi-pooling partitions is a crucial step in setting up a reduced optimization problem: First, one can solve for the optimal cluster centroids and their weights, i.e., the sum of data point shares assigned to each cluster. Exploiting the bi-pooling structure, the corresponding probabilistic partition can be uniquely determined. The minimization problem in the first step is considerably less complex

1. A point is an extreme point of a set if it cannot be represented as a convex combination of two points from that set.

than the original optimization problem because the number of variables and side constraints is substantially smaller—especially if the data set is large.

The classical k-means algorithm works as follows: Starting with k initial cluster centroids, each data point is assigned to the closest one. Based on this assignment, new centroids are formed. The process is repeated until an iteration is reached at which the centroids do not change anymore. By construction, this algorithm always converges to a monotone partition which is a local solution to the clustering problem. My derivations indicate that a k-means algorithm extended by the minimum distance constraint yields inefficient outcomes under certain circumstances: If the solution to the clustering problem is a bi-pooling partition, the algorithm never converges to the global optimizer.

Even though I consider the class of probabilistic clusters, this argument is also valid when considering deterministic partitions only: Given an optimal probabilistic bi-pooling partition, construct a deterministic partition by allocating each data point to the cluster which it belongs to with highest probability.[2] Clustering methods are usually applied to large data sets. The larger the number of data points, the closer the constructed deterministic partition to the optimal one.

A typical argument vindicating the use of monotone partitions is their simple form. However, bi-pooling partitions are also quite tractable: Basically, they correspond to almost monotone partitions that allow for mixing between pairs of clusters only. The strength of the bi-pooling result is thus that one does not need to forgo tractability when allowing for all possible partitions instead of monotone ones only.

The original distribution which the data are drawn from is usually unknown, but one might have a guess about it that needs to be verified. The typical approach to this problem is the implementation of a non-parametric hypothesis test based on the empirical distribution function of the data. These techniques yield robust results, but there is also a downside to that: First, constructing the empirical distribution function is an elaborate task if the data record is considerably large. Second, how should one proceed if the original data are not available and one has access to the clustered data only? I propose a new hypothesis test on the underlying distribution which the data are drawn from. The test problem is to evaluate whether this distribution is uniform or not. It is a Kolmogorov-Smirnov test which is performed on the basis of the clustered data: Instead of testing whether empirical distribution function generated by the data set equals the uniform distribution, I consider the testing problem whether the optimal distribution over cluster centroids implied by the data equals the distribution over centroids that would be the solution of the clustering problem if the underlying empirical distribution function could be the uniform distribution. The latter distribution over clusters is thus the solution to a continuous version of the constrained clustering problem, which is solved in Chapter 1.

2. In case of indifference, i.e., if a data point belongs to two clusters with the same probability, choose one of them at random.

Related Literature. This paper contributes to the literature on constrained cluster analysis (see Basu, Davidson, and Wagstaff (2008) or Grossi, Romei, and Turini (2017) for an excellent literature survey). The primary focus in this literature is on pairwise constraints such as must-link and cannot-link constraints. These conditions require that a pair of data points either must be assigned to the same cluster or cannot belong to the same one. Babaki, Guns, and Nijssen (2014) consider the class of antimonotone constraints that includes, among others, must-link and cannot-link constraints as well as constraints on the maximal number of data points per cluster. The minimum distance constraint analyzed in this paper is not antimonotone, and to the best of my knowledge, a constraint of this form hasn't been studied before in this literature. Perhaps closest related to the minimum distance constraint is the DP-means algorithm introduced by Kulis and Jordan (2012) or the λ-means algorithm by Comiter et al. (2016). The DP-means algorithm is an improved k-means algorithm, where new clusters are added at an iteration of the process if the distance of a data point to its closest cluster centroid is sufficiently large. A minimum distance evolves endogenously from running this algorithm, and depends on a prespecified model parameter. The λ-means algorithm is an extension of this process that allows estimating this parameter. This work differs from these two approaches because the minimum distance parameter is set exogenously.

The majority of research in this field uses an algorithmic approach (see, e.g., Kulis and Jordan (2012) or Comiter et al. (2016)) or optimization techniques from linear programming (see, e.g., Babaki, Guns, and Nijssen (2014) or Ganji, Bailey, and Stuckey (2016)). This paper uses techniques from the field of convex analysis and majorization (see Bertsekas, Nedic, and Ozdaglar (2003)).

This paper is also linked to the literature on information theory and quantization, which is the theory on approximating continuous-valued objects by a discretized one (for an excellent review of this strand of literature, let me refer to Gray and Neuhoff (1998)). The hypothetical distribution over clusters, that is used in this paper to perform the test for uniformity, can be interpreted as the output of a uniform quantization process. Recall that this distribution is a non-monotone bi-pooling partition in general. The standard approach in unconstrained quantization is a division of the continuous object into a discrete number of convex subsets (cf. Gray and Neuhoff (1998)). This technique extends to the work on constrained quantization embedding constraints on distortion, such as entropy-constraints (see Chou, Lookabaugh, and Gray (1989)), or on the shape of the discretized object such as the location or weights of its points (see Xu and Berger (2019)). Consequently, standard quantization techniques yield monotone partitions always, and can therefore not be applied to this paper's clustering problem. I am not aware of comparative work in the field of quantization that allows for this form of non-monotonicity. Amid monotone quantizers, the connection between quantization and majorization has been observed (see Baker (2015), for instance).

The paper is structured as follows: Section 3.2 formalizes the constrained clustering problem. Section 3.3 provides results on its analytic solution, and the reduced optimization problem. In Section 3.4, the hypothesis test for uniformity is proposed. Finally, Section 3.5 contains a discussion, and Section 3.6 concludes.

3.2 Constrained Clustering Problem

There is a set X of $n \in \mathbb{N}$ real-valued data points $x_1, \ldots, x_n$, arranged in non-decreasing order.

Definition 3.1. A *probabilistic partition* $P = (\mathscr{C}, w)$ consists of a finite set of clusters $\mathscr{C}$ and a weight function $w : X \times \mathscr{C} \to [0,1]$ so that

(i) $\sum_{C \in \mathscr{C}} w(x, C) = 1$ for all $x \in X$, and

(ii) $\sum_{x \in X} w(x, C) > 0$ for all $C \in \mathscr{C}$.

Probabilistic partitions assign shares $w(x, C)$ of data points x to clusters C. This generalizes the concept of a partition where each data point x is allocated to exactly one cluster C, that is, $w(x, C) \in \{0, 1\}$. The centroid of cluster C,

$$\mu_C = \frac{\sum_{x \in X} w(x, C) x}{\sum_{x \in X} w(x, C)}, \tag{3.1}$$

is the mean of the data points in that cluster weighted by their shares.

The aim is to find a probabilistic partition that minimizes the weighted average Euclidean distance between the data points and their corresponding cluster centroids subject to imposing a minimal distance between all cluster centroids. Formally, the objective is to solve

$$\min_{(\mathscr{C}, w)} \sum_{C \in \mathscr{C}} \sum_{x \in X} w(x, C) \left(x - \mu_{\mathscr{C}}\right)^2 \tag{3.2}$$

$$\text{s.t.} \quad | \mu_C - \mu_{C'} | \geq K \quad \text{for all } C, C' \in \mathscr{C}, \tag{3.3}$$

for some constant $K \geq 0$. This is a generalization of the basic probabilistic clustering problem (see Höppner et al. (1999)) without any side constraints ($K = 0$).

3.3 Solving the Model

3.3.1 Bi-Pooling Partitions

This section studies a specific class of probabilistic partitions:

Definition 3.2. A probabilistic partition P is a *monotone partition* if the following holds true for all $C \in \mathscr{C}$: For any $x, x' \in X$ with $x \leq x'$ such that $w(x, C), w(x', C) > 0$, it follows that $w(y, C) = 1$ for all $y \in X$ with $x < y < x'$.

Under a monotone partition, if a positive share of two data points is assigned to a cluster, then any data point in between is fully allocated to that cluster. The following figure shows a monotone partition. There are twelve data points, represented as

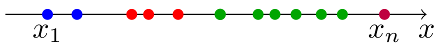


Figure 3.1. Example of a monotone partition

dots on the axis. Cluster membership is visualized by the dots color. There are four different clusters. The first two data points belong to the blue cluster, the next three ones to the red cluster, the last one to the purple cluster, and all remaining ones to the green cluster. Due to monotonicity, there is no overlap between the clusters meaning that all data points of one cluster are larger or smaller than all data points of another cluster.

Definition 3.3. A probabilistic partition P is a *bi-pooling partition* if the following holds true for all $C \in \mathscr{C}$: For any $x, x' \in X$ with $x \leq x'$ such that $w(x, C), w(x', C) > 0$, it follows that there exist some $C' \in \mathscr{C}$ such that $w(y, C) + w(y, C') = 1$ for all $y \in X$ with $x < y < x'$.

A cluster C is said to form a 2-*partition* with cluster $C' \neq C$ (and vice versa) if there exist $x, y, x' \in X$ with $x < y < x'$ so that $w(x, C), w(y, C'), w(x', C) > 0$. If C does not form a 2-partition with any other cluster C', it is said to form a 1-partition.

Under a bi-pooling partition, if a positive share of two data points is assigned to a cluster, then any data point in between is allocated to that cluster or to one other cluster with a positive share. Hence, there can be an overlap between different clusters, but only between two adjacent ones. So at most two clusters are pooled, therefore the name. The left graphic in Figure 3.2 depicts an example of a bi-pooling policy: The blue and purple cluster are completely separated, they form 1-partitions, whereas the red and green cluster are pooled and thus form a 2-partition. The right graphic illustrates a partition that is not bi-pooling: Three clusters, namely the red, green and purple one, are overlapping. By construction, any monotone partition is

(a) **(b)**

Figure 3.2. Example of a bi-pooling partition on the left side and a counterexample on the right side

a bi-pooling partition, but the converse does not hold true.

It turns out that bi-pooling partitions are optimal under the minimum distance constraint:

Theorem 3.4. *For any $K \geq 0$, the constrained clustering problem is solved by a bi-pooling partition.*

Proof. See Corollary 1.34. □

Bi-pooling partitions are the extreme points of the set of probabilistic partitions satisfying the minimum distance constraint (3.3). Linearity of the objective function (3.2) implies their optimality.

If the minimum distance constraint vanishes, i.e., if $K = 0$, an optimal partition can be found within the set of monotone partitions:

Corollary 3.5. *If $K = 0$, the constrained clustering problem is solved by a monotone partition.*

Proof. See the appendix. □

Intuitively, monotone partitions outperform bi-pooling partitions because data points are assigned to their closest cluster centroid. Without any side constraints, they can be implemented readily. In the presence of a minimum distance constraint, this is no longer true: Efficient monotone partitions might not fulfill the minimum distance constraints so that one resorts to bi-pooling partitions.

While monotone partitions are prevalent in the literature on cluster analysis, Theorem 3.4 suggests that it is with loss of generality to restrict attention to monotone partitions, as illustrated in the following example:

Example 3.6. Let $X = (0, 2, 3.5, 5.5, 7, 9)$, and suppose $K = 3$. The best monotone partition is depicted in the left graph of the figure below: There are three clusters, each consisting of two data points. The cluster centroids are 1, 4.5, and 8, which fulfill the minimum distance constraints. Can we find a better non-monotone partition? The answer is yes. The right graph illustrates the optimal probabilistic partition: There are four clusters. The blue and purple cluster consist of only one data point, the red cluster consists of data point 3.5 and share 0.9 (0.1) of data point 2 (7). Similarly, the green cluster consists of 5.5 and the remaining shares of 2 and 7. Notice that this is a bi-pooling partition. Moreover, the cluster centroids are 0, 3, 6, and 9, so the minimum distance constraint is fulfilled. Compared to the optimal

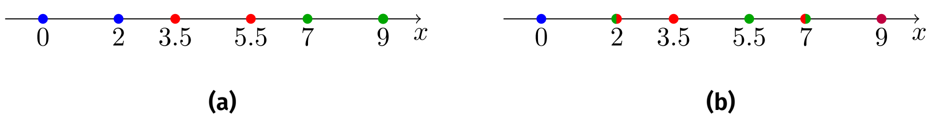


Figure 3.3. Optimal monotone partition on the left side and optimal bi-pooling partition on the right side

monotone partition, there is one additional cluster. This is why the bi-pooling partition outperforms the monotone partition. As argued above, monotone partitions are more efficient than bi-pooling partitions: In the optimal bi-pooling partition on the right graph, it would be better to assign data point 7 to the green cluster only, for instance. However, reshuffling shares to create a monotone partition with four clusters is infeasible. To see this, note that uncompressing the red and green cluster increases the distance between their centroids. This yields a failure of the minimum distance constraint between the blue and red cluster or between the green and the purple cluster. So the question is whether or not to add one more cluster at the expense of giving up on monotonicity. Here, this tradeoff is in favor of the additional cluster.

Besides, one can even argue that the bi-pooling partition is more reasonable by taking a closer look at the data points. The smallest and the largest data point are far apart from all other observations, indicating that each of them should form an individual cluster consisting only of this data point.

For tractability, the data set of this example is rather small, but the underlying idea is also applicable to larger data records.

A usual argument supporting the use of monotone partitions is their simple construction. Notwithstanding, the strength of Theorem 3.4 is that optimal probabilistic partitions take a relatively straightforward form. Consequently, a better partition can be reached by allowing for arbitrary probabilistic partitions without losing much tractability.

3.3.2 Simplified Problem

This section presents a simplified method on how to find an optimal bi-pooling partition given a positive minimum distance.

First, notice that the number of different clusters can be bounded by the minimum distance parameter $K > 0$: The smallest cluster centroid is at least x_1, and the largest cluster centroid does not exceed x_n. Thus, the distance between the smallest and the largest centroid must be no more than $x_n - x_1$. On the other hand, if there are m different cluster centroids, the distance between the smallest and largest centroid cannot exceed $(m-1)K$. Together, this yields

$$(m-1)K \leq x_n - x_1.$$

Consequently, the number of clusters in a probabilistic partition satisfying the minimum distance constraint (3.3) is at most $\frac{x_n - x_1}{K} + 1$.

The first step towards finding an optimal probabilistic partition with m clusters is to determine the optimal cluster centroids and their weights, i.e., the sum of data point shares associated with each cluster C, which is $\sum_{x \in X} w(x, C)$. Let F_n denote the empirical distribution function generated by the n data points, and let G_p be the distribution over cluster centroids implied by the probabilistic partition P. By construction, G_P is a mean-preserving contraction of F_n, that is, they generate equal means,

$$\int_{x_1}^{x_n} y\ dF_n(y) = \int_{x_1}^{x_n} y\ dG_P(y), \tag{3.4}$$

and F_n second-order stochastically dominates G_P,

$$\int_{x_1}^{x} F_n(y)\ dy \geq \int_{x_1}^{x} G_P(y)\ dy \ \text{ for all } x \in [x_1, x_n]. \tag{3.5}$$

The problem of finding the optimal cluster centroids and their weights is thus equivalent to determining the optimal mean-preserving contraction of F_n. Furthermore, the objective function (3.2) is equal to

$$\sum_{C\in\mathscr{C}}\sum_{x\in X} w(x,C)x^2 - 2x\mu_{\mathscr{C}} + \mu_{\mathscr{C}}^2 = \sum_{x\in X} x^2 - \sum_{C\in\mathscr{C}}\left(\sum_{x\in X} w(x,C)\right)\mu_{\mathscr{C}}^2, \tag{3.6}$$

where the first term on the right-hand side, $\sum_{x\in X} x^2$, is a constant, and can therefore be dropped from the optimization problem.

Let $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_m)$ with $\mu_1 < \ldots, < \mu_m$ be the m-dimensional vector containing the m realizations of a mean preserving contraction of F_n, and let $\boldsymbol{p} = (p_1, \ldots, p_m)$ be the vector containing the corresponding probabilities. The simplified optimization problem becomes

$$\max_{(\boldsymbol{\mu},\boldsymbol{p})\in[x_1,x_n]^m\times[0,1]^m} \sum_{i=1}^{m} p_i \cdot \mu_i^2 \tag{3.7}$$

$$\text{s.t.} \quad |\ \mu_i - \mu_{i-1}\ | \geq K \quad \text{for all } i \in \{2, \ldots, m\}, \tag{3.8}$$

$$\sum_{i=1}^{m} p_i = 1 \tag{3.9}$$

$$\sum_{i=1}^{m} p_i \mu_i = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3.10}$$

$$\text{for all } i \in \{1, \ldots, m\} : \sum_{k=1}^{j-1} \frac{1}{n}(x_j - x_k) \geq \sum_{k=1}^{i} p_k(x_j - \mu_k) \tag{3.11}$$
$$\text{for all } j \in \{1, \ldots, n\} \text{ with } x_j \in [\mu_i, \mu_{i+1}],$$

where $\mu_{m+1} \equiv x_n$. The objective function (3.8) corresponds to the negative of the original one (3.2) rescaled by the constant $\sum_{x\in X} x^2$. That is why the problem has become a maximization rather than a minimization problem. The inequality constraints (3.8) are the minimum distance constraints. (3.9) means that the probabilities of the different clusters add up to 1, and (3.4) ensures equal means as imposed by (3.10). The constraints (3.11) are sufficient to ensure second order stochastic dominance, that is, only a finite number of the constraints (3.5) needs to hold. To see this, recall that F_n and G_P are discrete distributions implying that the integrals $\int_{x_1}^{x} F_n(y)\,dy$ and $\int_{x_1}^{x} G_P(y)\,dy$ are piecewise linear and convex, as illustrated in Figure 3.4:

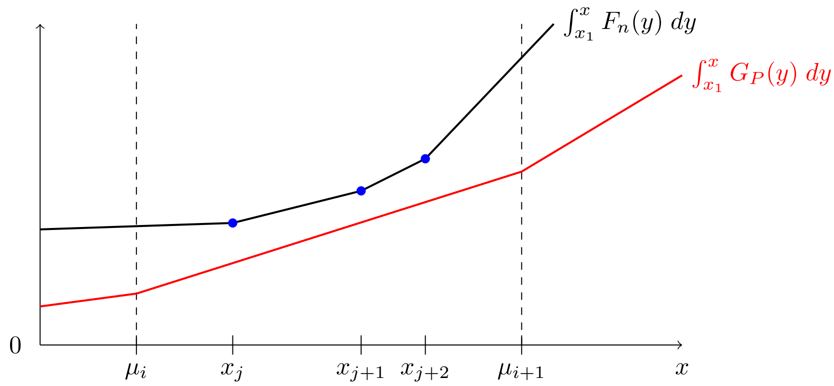


Figure 3.4. Section of the integral functions of F_n and its mean-preserving contraction G_P

The diagram shows the integral functions $\int_{x_1}^{x} F_n(y)\,dy$ and $\int_{x_1}^{x} G_P(y)\,dy$ on an interval between two centroids μ_i and μ_{i+1}. The graph of $\int_{x_1}^{x} G_P(y)\,dy$ is linear on that interval because G_P has no realization on its interior. The graph of $\int_{x_1}^{x} F_n(y)\,dy$ must lie above it by second order stochastic dominance. Piecewise linearity and convexity imply that this is guaranteed whenever this holds true at all its kink points, which are marked in blue.

Observe that it is easier to solve the above maximization problem than the original minimization problem on page 117, especially if there are many data points (n large), because there are considerably less optimization variables ($2m$ instead of nm) and also substantially less side constraints.

Using Theorem 3.4, the optimal probabilistic partition can immediately be derived from the optimal mean-preserving contraction of F_n. First, note that if $\int_{x_1}^{x} F_n(y)\,dy = \int_{x_1}^{x} G_P(y)\,dy$ for some $x \in [x_1, x_n]$, $w(y, C) = 0$ for all (y, C) with $y < x$ and $\mu_C > x$ as well as all (y, C) with $y > x$ and $\mu_C < x$. That is, all data points below (above) the threshold x are only allocated to cluster centroids below (above) and including that threshold. To see how to apply this, let's go back to the setting of Example 3.6:

Example 3.7. The empirical distribution function F_n has realizations X, each with probability $\frac{1}{6}$. The distribution over cluster centroids of the optimal bi-pooling policy, G_P, has realizations $\boldsymbol{\mu} = (0, 3, 6, 9)$ with probabilities $\boldsymbol{p} = \left(\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6}\right)$. The following figure illustrates the integral functions of F_n and G_P: At $x = 1$, the two integral

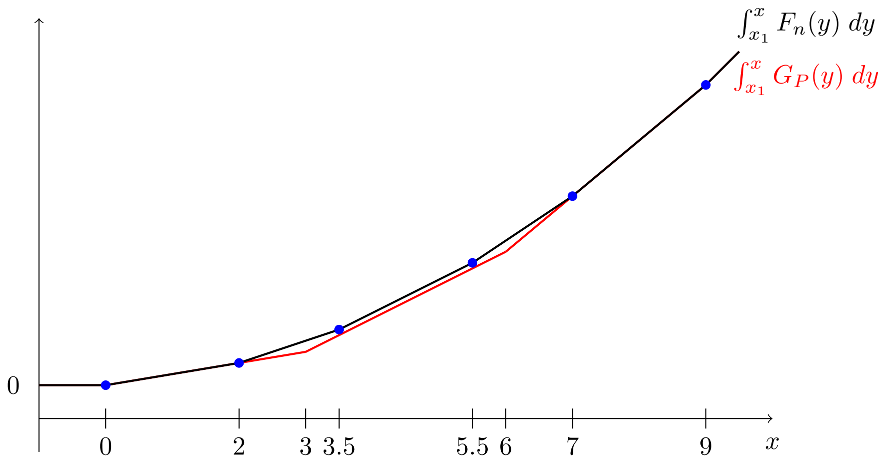


Figure 3.5. Integral functions of F_n and G_P

functions are equal. Consequently, $w(x, C) = 0$ for all $x < 1$ and all clusters C with $\mu_C > 1$, and $w(x, C) = 0$ for all $x > 1$ and all clusters with $\mu_C < 1$. Since 0 is the only data point and centroid below the threshold 1, the cluster with centroid 0 forms a 1-partition. Analogously, one can argue that the cluster with centroid 9 forms a 1-partition because the integral functions are equal at $x = 8$. For all $x \in (2, 7)$, the integral function of G_P is strictly smaller than the one of F_n. Consequently, the remaining four data points can be assigned to the remaining two clusters arbitrarily, subject to making sure that (3.1) holds for each such cluster and their weights equal the sum of assigned data point shares. There are several options to create such an assignment; one of them is the bi-pooling partition of Figure 3.3.

The observation from Example 3.7 that the first and last cluster form 1-partitions holds in general:

Proposition 3.8. *For any $K \geq 0$, the constrained clustering problem is solved by a bi-pooling partition such that the smallest and largest centroid belong to a 1-partition, respectively.*

Proof. See Lemma 1.38. □

This finding allows to draw further conclusions about the structure of an optimal bi-pooling policy if the number of clusters is small: If there are only two or three clusters, there is an optimal bi-pooling policy. In particular, for the case of three clusters, observe that if the first and the last cluster form a 1-partition, then the intermediate one has to form a 1-partition as well.

To conclude, an optimal bi-pooling partition can be determined using the following three-step procedure:

Step 1: For each $m \in \mathbb{N} \cap \left[1, \frac{x_n - x_1}{K} + 1\right]$, find the optimal mean-preserving contraction G_m^* of the empirical distribution function F_n (if it exists)[3].

Step 2: Compare the mean-preserving contractions $G_1^*, \ldots, G_m^*$ and select the best one out of them.

Step 3: Determine the optimal bi-pooling partition based on the optimal mean-preserving contraction.

3.4 Test for Uniformity

The original distribution F of the data is unknown, but one often has a guess about it. One way to verify this guess is to perform a Kolmogorov-Smirnov test based on the empirical distribution function F_n. If the set of data points is very large, construction of the empirical distribution function can be quite time-consuming. Beyond that, what is the appropriate procedure if the original data are not available, but only the clustered data?

This section proposes a test for the hypothesis that the data points are drawn from a uniform distribution on $[0, 1]$ that does not rely on the data points itself, but only on the clustered information:

$$H_0 \; : \; F = \text{Uni}[0,1] \quad \text{versus} \quad H_1 \; : \; F \neq \text{Uni}[0,1]. \tag{3.12}$$

Let G_n be the optimal mean-preserving contraction of F_n, and let G^* be the optimal mean-preserving contraction of Uni$[0,1]$. This function G^* has been derived in Section 1.7.

Instead of the original hypotheses (3.12), consider the problem

$$H_0 \; : \; G = G^* \quad \text{versus} \quad H_1 \; : \; G \neq G^*. \tag{3.13}$$

To test for uniformity, one can run a Kolmogorov-Smirnov test with the usual test statistic $\sup_{x \in [0,1]} |G_n(x) - G^*(x)|$.

3.5 Discussions

3.5.1 Implications on *k*-Means Clustering

The basic k-means clustering problem aims for optimal deterministic partitions of the data set meaning that $w(x, C) \in \{0, 1\}$ for all data points x and all clusters C.

3. If G_m^* exists, then $G_{m'}^*$ exists for all $m' \leq m$. Moreover G_1^* always exists.

In this setting, monotone partitions are optimal.[4] Therefore, the k-means algorithm, whose outcome is a deterministic, monotone partition, is an appropriate method to find a numerical, approximate solution.

Introducing a minimum distance constraint annihilates this result. Theorem 3.4 suggests that the optimal deterministic partition can be a bi-pooling partition that is not a monotone partition (cf. Example 3.6). So my results suggest that applying the classical k-means algorithm in an environment with minimum distance constraints may lead to imprecise numerical solutions as it never converges to a nonmonotone partition.

3.5.2 Number of Clusters

Finding the right number of clusters for a data set is still an open research problem in the field of cluster analysis.

As shown in Section 3.3.2, adding a minimum distance K resolves this problem because the optimal number of clusters can be determined from the solution of the constrained clustering problem. Consequently, the minimum distance parameter K can be interpreted as a measure for the optimal number of clusters. This can be helpful for applications because it is typically easier to argue what is the right distance between clusters to ensure meaningful disparity among them—even without knowing detailed properties of the data set X. Otherwise, a certain amount of information, such as a diagram showing the data points' location, is necessary to make a reasonable statement on why a certain number of clusters is appropriate for the given data.

3.5.3 Extensions

There are several questions for future research that seem promising:

This paper studies one-dimensional data. It would be interesting to extend the results of the constrained clustering problem to the multidimensional case. While the representation of a mean-preserving contraction as depicted in Section 3.3.2 cannot be straightforwardly applied to spaces with more than one dimension, the concept of bi-pooling partitions can be generalized.

As argued in Section 3.5.1, a version of the typical k-means algorithm, supplemented by the minimum distance constraint, is not an adequate method to find numerical approximations. Hence, constructing an algorithm whose outcomes can be bi-pooling partitions, not only monotone partitions, would be the next step for future research.

Finally, the idea of the test for uniformity in Section 3.4 can be extended to initial distributions F_0 other than the uniform distribution on the unit interval. For

4. This follows from analogous arguments as those used in Corollary 3.5.

that, one needs to determine the optimal mean preserving contraction of any such F_0.

3.6 Conclusion

Summing up, this paper explores a probabilistic clustering problem endowed with a minimum distance constraint ensuring dissimilarity among clusters. The main qualitative result is that monotone partitions are not optimal in general. Typical clustering algorithms converge to monotone partitions, implying a need for an improved algorithm for the constrained setting. Besides, specifying the minimum distance constraint replaces the problem of finding the optimal number of clusters. Moreover, a test for uniformity of the initial distribution from which the data are drawn is suggested that can be performed based on the clustered information only; the original data are not needed for that.

Appendix 3.A

Proof of Corollary 3.5. Suppose to the contrary, the clustering problem is solved by a bi-pooling partition $P = (\mathscr{C}, w)$ that is not monotone. Consequently, there exist two different clusters C_1 and C_2 satisfying $w(x, C_1), w(y, C_2), w(x', C_1) > 0$ for some $x < y < x'$. Without loss of generality, suppose $\mu_{C_1} < \mu_{C_2}$.[5]

Construct a new probabilistic partition $P^* = (\mathscr{C}, w^*)$ with

$$\sum_{x\in X} w^*(x, C) = \sum_{x\in X} w(x, C) \text{ for all } C \in \{C_1, C_2\}, \tag{3.A.1}$$

$$w^*(x, C_1) = 0 \text{ for all } x > x^* \text{ and } w^*(x, C_2) = 0 \text{ for all } x < x^* \text{ for some } x^*, \tag{3.A.2}$$

$$w^*(x, C) = w(x, C) \text{ for all } x \in X \text{ and all } C \in \mathscr{C}\backslash\{C_1, C_2\}. \tag{3.A.3}$$

In words, the two clusters C_1 and C_2 are separated under the new partition by reassigning their data point shares among one another while keeping their weights equal: Data points falling below some cutoff x^* are only assigned to C_1 while data points exceeding this cutoff are allocated to C_2 only. By construction, it holds for the new centroids of the two clusters C_1 and C_2, denoted by $\mu^*_{C_1}$ and $\mu^*_{C_2}$, that $\mu^*_{C_1} < \mu_{C_1}$, $\mu^*_{C_2} > \mu_{C_2}$ and

$$\sum_{C\in\{C_1,C_2\}} \sum_{x\in X} w^*(x, C)\mu^*_C = \sum_{C\in\{C_1,C_2\}} \sum_{x\in X} w(x, C)\mu_C. \tag{3.A.4}$$

5. If not, switch the roles of C_1 and C_2.

Finally, observe that

$$\begin{aligned}
&\sum_{C\in\mathscr{C}}\sum_{x\in X} w^*(x,C)\left(x-\mu_{\mathscr{C}}\right)^2-\sum_{C\in\mathscr{C}}\sum_{x\in X} w(x,C)\left(x-\mu_{\mathscr{C}}\right)^2\\
&\quad=\sum_{x\in X}x^2-\sum_{C\in\mathscr{C}}\left(\sum_{x\in X}w^*(x,C)\right)\mu_{\mathscr{C}}^2-\left(\sum_{x\in X}x^2-\sum_{C\in\mathscr{C}}\left(\sum_{x\in X}w(x,C)\right)\mu_{\mathscr{C}}^2\right)\\
&\quad=-\sum_{C\in\{C_1,C_2\}}\left(\sum_{x\in X}w^*(x,C)\right)\left(\mu_{\mathscr{C}}^*\right)^2+\sum_{C\in\mathscr{C}}\left(\sum_{x\in X}w(x,C)\right)\mu_{\mathscr{C}}^2\\
&\quad<\ 0
\end{aligned}$$

where the first equality follows from (3.6) and the second one from (3.A.3). The inequality is due to (3.A.1), (3.A.4) and the fact that $f(x)=x^2$ is a convex function.

References

Babaki, Behrouz, Tias Guns, and Siegfried Nijssen. 2014. "Constrained Clustering Using Column Generation." In *International Conference on Integration of Constraint Programming, Artificial Intelligence, and Operations Research*, 438–54. Springer. [116]

Baker, David M. 2015. "Quantizations of probability measures and preservation of the convex order." *Statistics & Probability Letters* 107: 280–85. [116]

Basu, Sugato, Ian Davidson, and Kiri Wagstaff. 2008. *Constrained Clustering: Advances in Algorithms, Theory, and Applications.* CRC Press. [116]

Bertsekas, Dimitri, Angelia Nedic, and Asuman Ozdaglar. 2003. *Convex Analysis and Optimization.* Vol. 1. Athena Scientific. [114, 116]

Chou, Philip A, Tom Lookabaugh, and Robert M Gray. 1989. "Entropy-constrained vector quantization." *IEEE Transactions on acoustics, speech, and signal processing* 37 (1): 31–42. [116]

Comiter, Marcus, Miriam Cha, HT Kung, and Surat Teerapittayanon. 2016. "Lambda Means Clustering: Automatic Parameter Search and Distributed Computing Implementation." In *2016 23rd international conference on pattern recognition (ICPR)*, 2331–37. IEEE. [116]

Ganji, Mohadeseh, James Bailey, and Peter J Stuckey. 2016. "Lagrangian Constrained Clustering." In *Proceedings of the 2016 SIAM International Conference on Data Mining*, 288–96. SIAM. [116]

Gray, Robert M., and David L. Neuhoff. 1998. "Quantization." *IEEE transactions on information theory* 44 (6): 2325–83. [116]

Grossi, Valerio, Andrea Romei, and Franco Turini. 2017. "Survey on using constraints in data mining." *Data mining and knowledge discovery* 31 (2): 424–64. [116]

Höppner, Frank, Frank Klawonn, Rudolf Kruse, and Thomas Runkler. 1999. *Fuzzy cluster analysis: methods for classification, data analysis and image recognition.* John Wiley & Sons. [117]

Kulis, Brian, and Michael I. Jordan. 2012. "Revisiting K-Means: New Algorithms via Bayesian Nonparametrics." In *Proceedings of the 29th International Conference on Machine Learning*, 1131–38. ICML'12. Omnipress. [116]

Xu, Chuang, and Arno Berger. 2019. "Best finite constrained approximations of one-dimensional probabilities." *Journal of approximation theory* 244: 1–36. [116]