NORTHWESTERN UNIVERSITY

Essays in Applied Microeconomics

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Managerial Economics and Strategy

By

Philip Marx

EVANSTON, ILLINOIS

December 2017

ABSTRACT

Essays in Applied Microeconomics

Philip Marx

This dissertation consists of three essays in applied microeconomics. In the first chapter, I introduce a new statistical test for identifying prejudice from empirical data. In the second chapter, I (joint with James Schummer) consider the revenue maximization problem for a two-sided, one-to-one matching platform. In the third chapter, I (joint with Tomer Hamami) consider the dynamic incentives of news firms to polarize consumers in order to increase rent extraction.

Disparities along racial and ethnic lines persist across domains, and distinguishing among the possible sources of such disparities matters. Chapter 1 introduces the first *absolute* test for identifying prejudice in the common setting where the returns to treatment intensity are weakly diminishing for each treator. As examples, law enforcement officers engage in stops and searches, employers hire, judges deny bail, doctors administer procedures and medical tests, and creditors extend loans and mortgages. The test additionally unifies the existing literature and can be partially extended to the important setting where treators vary in the quality of their information. Empirically, the test finds

evidence of prejudice in a dataset where the existing test does not. Methodologically, the test is the first in the literature to jointly use actions *and* outcomes in the test statistic.

Chapter 2 considers the revenue maximization problem for one-to-one matching platforms that choose among stable matching mechanisms. In this sense the work bridges an existing gap between two literatures in economics: matching and two-sided markets. In seeming contrast with recent results in the matching literature, a platform does *not* want to price discriminate across the two sides of the market based on their relative sizes when preferences are independent. The analysis introduces a new class of matching mechanisms and an approximation for the stable platform's expected revenue using an exact expression for revenue under a *constrained serial dictatorship* mechanism. The chapter concludes by considering how correlation in agents' preferences affects results and how this contrasts with classic models of pricing in two-sided markets.

Chapter 3 addresses a gap in the literature on political slant in news media regarding the means and incentives news firms have to affect their consumers' beliefs over time. Specifically, competitive tensions generate an incentive for media firms to intentionally polarize consumers in order to increase product differentiation, and this can occur even if consumers are initially homogeneous and unbiased. Computational analysis and simulations corroborate this analysis.

Acknowledgements

I am extremely grateful to my dissertation committee – Nicola Persico, James Schummer, Willemien Kets, and Jonathan Guryan – for their wisdom, advice, encouragement, patience, and support throughout the Ph.D. process. I am additionally grateful to David Besanko, Nicola Bianchi, Tomer Hamami, Alexander Limonov, Charles Manski, Daniel Martin, Robert Porter, Colin Shopp, Amanda Starc, Yuta Takahashi, Rajkamal Vasu and participants at seminars at Northwestern and the 2016 GAMES conference for helpful discussions and feedback.

Dedication

To my parents and to my siblings – Raphael, Maurice, and Inna – for their unconditional love and unwavering support.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

An Absolute Test of Racial Prejudice

1.1. Introduction

Racial and ethnic disparities persist in the criminal justice system, law enforcement, education, employment, health care, and housing (NRC, 2004).Two sources of such disparities – *statistical discrimination* and *prejudice* – are identified in the literature on the economics of discrimination. To illustrate each concept, consider the problem of a *treator* who makes decisions under uncertainty over a set of *treatable* agents. The optimal treatment decision depends on an imperfectly observed *quality* of the treatable agent. For example, a police officer (treator) chooses which drivers (treatable) to search (treat) for contraband (quality); a doctor chooses which patients to test for a medical condition.[1] Disparities in treatment can arise along any treatable characteristic such as race, gender, sexual orientation, or provenance. For clarity, however, the discussion proceeds with race. Racial disparities are driven by *prejudice* if the treator has a *utility* motive for treating agents differently.[2] In this case the treator is called *prejudiced*. Alternatively, racial disparities are driven by *statistical discrimination* if the treator has an *informational* motive.[3] This might happen when race correlates with the unobserved quality.

[1] Other examples include a judge who seeks to convict the guilty or an employer who seeks to hire productive workers.
[2] Following Becker (1957).
[3] Following Arrow (1973) and Phelps (1972).

Distinguishing between prejudice and statistical discrimination matters. For the policy-maker who seeks to mitigate existing disparities, prejudice among treators calls for different measures than an undesirable equilibrium outcome. For example, a temporary policy of affirmative action can resolve the issue of equilibrium selection [e.g. Coate and Loury (1993)], but not prejudice.[4] This underscores the value of identifying prejudice in the presence of statistical discrimination.

However, empirically distinguishing between sources of disparity is challenging. In particular, regression analysis in this setting places stringent demands on the data available to the researcher. *Action*-based regression tests that explain (differences in) actions as a function of characteristics observable to the researcher are susceptible to omitted variables bias if the researcher does not observe all of the information used by the treator; conversely, such tests may fail to find discrimination if treators, in order to enshroud their prejudice, condition decisions only on characteristics correlated with race. Even with evidence that treatment depends on race, such regression-based action tests do not distinguish whether disparities stem from prejudice or from statistical discrimination.

Alternatively, *outcome*-based tests compare the average quality of treated agents across race. This idea originates with Becker (1957) and is first operationalized by Knowles, Persico, and Todd (2001) [KPT]. KPT show that a comparison of average outcomes across treatable race provides a robust test for prejudice when treators face *constant*

[4] The respective concepts also correspond loosely to the claims of disparate treatment and disparate impact in anti-discrimination law. A treator engages in disparate treatment when he or she treats similarly situated individuals differently because of a protected characteristic. A treator engages in disparate impact when decisions are made according to a formally neutral policy that adversely affects a protected group of people. While disparate treatment is always unlawful, the legality of disparate impact depends on the domain and on the justification for the policy in question.

marginal returns to *individual* treatment intensity.[5] In the context of policing, an individual officer choosing how often to search a motorist group perceives the expected marginal return of searching the first motorist to be the same as the expected marginal return of searching the last motorist. This arises for two reasons. First, given a fixed treatment intensity, the officer does not perceive any differences in the likelihood of guilt among members of the motorist group. Second, the *individual* choice of treatment intensity has no impact on the motorist decision to carry contraband.[6]

The key assumption of constant marginal returns to individual treatment intensity need not hold, however. In policing, motorist characteristics observed by the officer may correlate with guilt; in medical care, patient symptoms may correlate with the incidence of a medical condition. Both examples violate the first condition of the previous paragraph. For a given treator, the marginal return to the first unit of treatment intensity – for example, searching only the guiltiest drivers or testing only the highest-risk patients – is higher than the marginal return of the last unit of treatment intensity. Therefore the individual treator faces *diminishing* marginal returns to treatment intensity.[7]

Anwar and Fang (2006) [AF] introduce a test for prejudice in the case of diminishing returns to individual treatment intensity. The test relies on an observable source of variation in treatment intensity across treators for a fixed group of treatable agents. However, the test only identifies prejudice in a partial sense. Namely, the AF test is

[5] An earlier paper by Ayres and Waldfogel (1994) proposes a closely related test based on constant returns to treatment without explicitly resolving why returns should be constant.
[6] In contrast, marginal returns to *aggregate* treatment intensity must incorporate any effects of increased treatment intensity on the behavior of treatable agents – for example, motorist decisions to carry contraband or job applicant incentives to invest in human capital.
[7] Diminishing marginal returns to treatment intensity give rise to infra-marginally treated agents, which again complicates the task of identifying prejudice. This is referred to as the infra-marginality problem in the racial profiling literature.

relative: it cannot disentangle competing hypotheses of who is prejudiced against whom. In contrast, a test is defined to be *absolute* if it can identify a specific prejudiced treator and a specific group of victimized treatable agents. To contextualize the improvement, consider the example of racial profiling: only an absolute test can distinguish whether minorities are victims or beneficiaries of current policing practices.

This paper presents the first absolute test for prejudice when marginal returns to individual treatment intensity may be diminishing. The test requires no additional theoretical assumptions over the existing literature and relies on the same source of identification – variation in behavior across treators – as AF. Therefore the improvement in explanatory power stems from an additional contribution: the test is the first to jointly use actions *and* outcomes in the test statistic.

The key insight of the test is that variation in treator behavior gives rise to a group of "inter-marginal" agents in each treatable group. For example, consider two police officers who search the same treatable group of motorists with different intensities. This partitions the searched treatable motorists into two groups: those who are searched by both officers, and those who are searched only by the high-intensity officer. The latter group is "inter-marginal:" the expected return of searching any member of the group is weakly higher than the search cost of the high-intensity officer but weakly lower than the search cost of the low-intensity officer.[8] Therefore the average return to searching an inter-marginal agent bounds the treatment costs. Discrepancies in treatment costs across treatable groups in turn define prejudice. Finally, partially identified treatment costs provide a test for prejudice: if the sets of possible treatment costs by treatable group do

[8] Otherwise officer behavior is suboptimal.

not overlap for a given treator, the treator must be prejudiced. Implementing the test requires only one source of variation across each group, but further granularity in treator and treatable characteristics enhances explanatory power.

The new model and test also unify the existing literature. The model extends the AF model to encompass the KPT model by permitting constant returns to individual treatment intensity. A generalized model in turn allows for a comparison of tests. When the data exhibit diminishing returns to individual treatment intensity, the new test is strictly more powerful than the AF test: empirical evidence of prejudice under the AF test also constitutes evidence of prejudice under the new test, but not vice versa. Insofar as the the data exhibit constant returns to individual treatment intensity, the test statistic coincides with the average outcome of KPT.

An extension of the model considers the case where signals vary across treators. The existing assumption of equal information across treators is limiting.[9] For example, the assumption prohibits using tenure (or any correlated variable) as a source of treator variation without excluding the possibility that treators learn with experience.[10] Many scenarios with differential signals are captured by differences in signal *informativeness*. Extending the model to allow for differences in signal informativeness across treators yields an analogous, albeit weakened, test for prejudice.

To demonstrate the empirical power of the new test, I revisit the application of AF. The new test finds the *first* evidence of prejudice in the AF dataset on Florida State

[9] This assumption is also made by existing tests [KPT, AF].
[10] This also rules out applying the test in law enforcement if same-race policing confers informational advantages [as suggested for local policing by Donohue and Levitt (2001)].

Highway Patrol stops and searches. Namely, Hispanic officers exhibit prejudice against Hispanic motorists.[11]

The paper proceeds as follows. Section 1.2 provides additional discussion of the related literature. Section 1.3 presents the model of treator behavior and defines the researcher's problem. Section 1.4 presents the empirical test. Section 1.5 extends the model and test to the case where the informativeness of signals may vary across treators. Section 1.6 provides an empirical application on the AF data. Section 1.7 concludes.

1.2. Related Literature

The test belongs to a literature on racial profiling that derives testable implications of prejudice in terms of simple summary statistics of the data.[12] In a seminal paper, Knowles, Persico, and Todd (2001) show that a comparison of average outcomes provides a test for prejudice when the marginal returns to treating an additional agent are constant for treators. This is a result of equilibrium in their model: if treatable agents best-respond to intensity of treatment, a no-arbitrage condition implies that average outcomes should equal treatment costs.[13]

However, Anwar and Fang (2006) provide evidence that the marginal returns to individual treatment intensity may not be constant. In this setting they exploit variation in treator behavior to provide a new test for prejudice. Similarly, Antonovics and Knight

[11] In addition, a new test for informativeness rejects the assumption that officers are equally informed. In particular, white and Hispanic officers are not equally informed over Hispanic motorists. Nevertheless the finding of prejudice under the absolute test is robust to a weakened set of assumptions on informativeness across officers that is consistent with the data.

[12] A large literature in labor also attempts to distinguish between statistical discrimination and prejudice. See Charles and Guryan (2011) for an overview and Altonji and Pierret (2001) for an example that tests for race-based statistical discrimination.

[13] See also Persico and Todd (2006) for an additional class of decision problems for which the outcome test is valid.

(2009) observe that treators may vary in their behavior and use this to develop a related but parametric test for prejudice. Both papers precede the current one in their reliance on variation in treator behavior to identify prejudice.

The new test for prejudice relies on *partially identifying* the parameters that underlie prejudice. In this sense the test is related to a large literature on partial identification (see Manski (2003) and Tamer (2010) for overviews). In the context of testing for prejudice in policing, Hernandez-Murillo and Knowles (2004) also derive a test using bounds. However, their bounds stem from the researcher's uncertainty over the proportion of searches that are at the discretion of the decision-maker.

The test assumes, like KPT and AF, that the treator's objective function is to maximize outcomes net of the cost of treatment. Other objective functions may guide decision-making, in which case the test does not generally apply. For example, in the context of policing, Dominitz and Knowles (2005) and Manski (2006) consider problems where the objective is to minimize the crime rate and a social cost function,[14] respectively.

1.3. The Model

This section develops the model that underlies the empirical test presented in Section 1.4. It begins with the treator's decision problem and continues with the researcher's problem of identifying prejudice.

[14] The social cost function includes the costs of crime, punishment, and search.

1.3.1. The Treator's Decision Problem

A continuum of treators of type t or t' decide whether to treat an infinite and independent stream of treatable agents of race $r \in \{M, W\}$.[15] The decision to treat, or the *action*, is denoted by $a \in \{0, 1\}$. The quality of a treatable agent is denoted by q and can belong to any real-valued set $\mathcal{Q}$. To simplify notation I assume that $\mathcal{Q} \subset \mathbb{R}_+$. Quality is only revealed when an agent is treated. In this case I refer to revealed quality as the *outcome*.

The payoff to a treator of race i who takes action a on a treatable agent of race r and quality q is given by:

$$a \cdot [q - c_t(r)]. \tag{1.1}$$

A treator is *prejudiced* when the cost of treatment differs across treatable groups:

Definition 1.1. *Treator t is* prejudiced *if the cost of treatment depends on the treatable race:*

$$c_t(M) \neq c_t(W). \tag{1.2}$$

Imperfect information over quality introduces uncertainty into the treator's decision problem.[16] As in AF, I model this uncertainty through a single-dimensional signal x which is correlated with q. The signal incorporates any observable characteristics of the treatable agents that correlate with quality. This includes any characteristics of the interaction between the treator and treatable agent.[17] Formally, let Q and X denote the random

[15] With slight abuse of notation I also refer to an agent in a group by the group identity, e.g. treator t and treated agent r.
[16] Otherwise statistical discrimination is precluded and inference over prejudice is straightforward.
[17] In policing, for example, a motorist may respond nervously to questioning, especially when guilty.

variables with respective realizations q, x. Let $F^r(q, x)$ denote the joint distribution of quality q and signal outcome x for treatable agents of race r. I assume that each family $\{F^r_Q(\cdot|x)\}_x$ has a complete FOSD order in x.[18]

A treator t chooses an optimal action given a signal outcome x for a treatable agent of race r:

$$a^*_t(x; r) \in \arg\max_a \, a \cdot [\mathbb{E}[Q_r|X_r = x] - c_t(r)]. \tag{1.3}$$

The optimal action is to treat an agent iff the expected benefit of treatment exceeds the cost $c_t(r)$. The assumptions on the signal imply the existence of an optimal cutoff rule.

Lemma 1.1 (AF, Proposition 1)**.** *Treator t has an optimal cutoff decision rule such that $a^*_t(x; r) = 1$ iff $x \geq x^*_r$.*

Proof: The result follows immediately from Lemma 1 of Athey and Levin (2017) because the signal satisfies FOSD and the utility function (1.1) is trivially supermodular in (a, q). □

Finally, note that the assumption that F is independent of the treator t may be unduly restrictive. In Section 1.5, I consider an extension where signal informativeness may vary across treator and treatable pairs.

---

[18] For simplicity I assume that there are no mass points on signals. This is without loss of generality. Namely, replace any signal outcome with positive mass by a continuum of signal outcomes that differ in name but generate the same posterior beliefs.

1.3.2. The Researcher's Problem

I now turn to the researcher's problem of inferring prejudice. The researcher observes *average* actions

$$\bar{a}_t(r) = \mathbb{E}[a_t^*(X_r; r)] \tag{1.4}$$

and *average* outcomes

$$\bar{q}_t(r) = \mathbb{E}[Q_r | a_t^*(X_r; r) = 1] \tag{1.5}$$

for each treator t and treatable group r.[19] Let $\bar{\mathbf{a}} = \big(\bar{a}_t(r)\big)_{rt}$ and $\bar{\mathbf{q}} = \big(\bar{q}_t(r)\big)_{rt}$ denote the vectors of average actions and average outcomes by treatable group and treator.

Inferring prejudice from disparities is complicated by the possibility of *statistical discrimination*: even the average actions of an unprejudiced treator may vary across treatable groups. Adopting the definition of KPT,

Definition 1.2. *Suppose that costs of treatment for treator i are equal across treatable groups: $c_t(M) = c_t(W)$. Then treator i exhibits statistical discrimination if the average action varies by treatable group: $\bar{a}_t(M) \neq \bar{a}_t(W)$.*

Two mechanisms underlie statistical discrimination in the model. The first, in the spirit of Arrow (1973), is that the distributions of quality can vary by treatable race:

$$Q_M \overset{d}{\neq} Q_W.$$

[19] The researcher may observe additional characteristics of the treatable groups. This can be incorporated by considering a finer r.

To illustrate in the context of policing, officers will search a motorist group more if the motorist group carries contraband at a higher rate, *ceteris paribus*. The second, in the spirit of Phelps (1972), is that treator information varies across treatable groups, even conditional on treatable quality:

$$F_X^M(\cdot|q) \neq F_X^W(\cdot|q).$$

This happens if officers search motorist groups at different rates because the officers' ability to distinguish between innocence and guilt varies across motorist groups.

To identify prejudice, the researcher derives testable implications of prejudice as a function of the observable vectors of average actions $\bar{\mathbf{a}}$ and outcomes $\bar{\mathbf{q}}$. The result is a *test* $T(\bar{\mathbf{a}}, \bar{\mathbf{q}})$.[20] Using vectors of observables allows the researcher to exploit variation in average actions across treators for a fixed group of treatable agents.[21]

Tests vary in the precision of the null hypotheses that they can reject. A test for prejudice should at least be able to identify prejudice – or equivalently, to reject a null of *no prejudice*. Adopting the terminology of AF, such a test is called *relative*.

Definition 1.3. *A test T_R is* relative *if it can reject a null hypothesis that no group of treators is prejudiced against any treatable group:*

$$\mathcal{H}_0 : c_t(M) = c_t(W) \quad \textit{for all treator types } t \textit{ and treatable groups } M, W.$$

The name stems from the relative nature of the alternate hypothesis: *some* group of treators is prejudiced against *some* group of treatable agents. A relative test can identify

[20] For simplicity of notation I omit the arguments to a test where possible.
[21] This insight, as well as the introduction of a source of variation, are important contributions of AF.

that prejudice exists, but it cannot disentangle finer hypotheses of which treator is prejudiced against which treatable group. A more specific test is an *absolute* one. An absolute test can identify specific perpetrators and victims of prejudice:

Definition 1.4. *A test T_A is* absolute *if it can reject a null hypothesis that treator t is not prejudiced against a* given *treatable group, say M:*

$$\mathcal{H}_0^{tMW} : c_t(M) \geq c_t(W)$$

The alternate hypothesis of an absolute test can identify a specific group of discriminating treators and a treatable group that is discriminated against. Note that an absolute test is also relative, but a relative test need not be absolute. Thus, absoluteness is a desirable property of a test.

1.4. An Absolute Test of Prejudice

The main theoretical contribution of the paper is to provide the first *absolute* test for prejudice when individual returns to treatment are diminishing.

Theorem 1.1. *Suppose individual returns to treatment are weakly diminishing. There exists an* absolute *test $T_A(\bar{\mathbf{a}}, \bar{\mathbf{q}})$ for prejudice based on partial identification of treatment costs:*

$$c_t(r) \in S^{r,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}}).^{22}$$

[22] Note that each set $S^{r,t}(\cdot)$ only relies on the subset of average actions and outcomes for treatable group r, which I omit for simplicity of notation.

Treator t is prejudiced if:

$$S^{M,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}}) \cap S^{W,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}}) = \emptyset. \tag{1.6}$$

Proof: The model partially identifies treatment costs when treators vary in their average actions over a fixed treatable group. Therefore suppose that treators t and t' take average actions $\bar{a}_t(r) > \bar{a}_{t'}(r)$ over treatable group r.

A set $\mathcal{X}_r^{high}$ of the highest signals generates a sufficiently high expected outcome for both treators to optimally treat. A set $\mathcal{X}_r^{int}$ of intermediate signals generates a sufficiently low expected outcome that treator t optimally treats and t' optimally does not treat.

The set of signals $\mathcal{X}_r^{high}$ is infra-marginal. Each signal in $\mathcal{X}_r^{high}$ generates a weakly higher expected outcome than the treatment cost of treator t'. The set of signals $\mathcal{X}_r^{int}$ is "inter-marginal:" any signal in $\mathcal{X}_r^{int}$ generates a weakly higher expected outcome than the treatment cost of treator t, and a weakly lower expected outcome than the treatment cost of treator t'. Otherwise the treatment decisions are suboptimal. Therefore the average outcome conditional on either set of signals places bounds on the the treatment costs of both treators. In particular, when $\bar{a}_t(r) > \bar{a}_{t'}(r)$,

$$E[Q_r | X_r \in \mathcal{X}_r^{high}] \geq c_{t'}(r) \geq E[Q_r | X_r \in \mathcal{X}_r^{int}] \geq c_t(r). \tag{1.7}$$

The next part of the identification argument expresses the conditional average outcomes in (1.7) as a function of the observables $\bar{\mathbf{a}}$ and $\bar{\mathbf{q}}$. The lower-intensity treator (by assumption t') treats if and only if the signal belongs to $\mathcal{X}_r^{high}$. It follows that:

$$\mathbb{E}[Q_r | X_r \in \mathcal{X}_r^{high}] = \bar{q}_{t'}(r). \tag{1.8}$$

For $\mathbb{E}[Q_r|X_r \in \mathcal{X}_r^{int}]$, the Law of Iterated Expectations implies:

$$
\begin{aligned}
\bar{q}_t(r) &= \mathbb{E}[Q_r|X_r \in \mathcal{X}_r^{high} \cup \mathcal{X}_r^{int}] \\
&= \frac{\bar{a}_{t'}(r)}{\bar{a}_t(r)}\mathbb{E}[Q_r|X_r \in \mathcal{X}_r^{high}] + \frac{\bar{a}_t(r) - \bar{a}_{t'}(r)}{\bar{a}_t(r)}\mathbb{E}[Q_r|X_r \in \mathcal{X}_r^{int}] \\
&= \frac{\bar{a}_{t'}(r)}{\bar{a}_t(r)}\bar{q}_{t'}(r) + \frac{\bar{a}_t(r) - \bar{a}_{t'}(r)}{\bar{a}_t(r)}\mathbb{E}[Q_r|X_r \in \mathcal{X}_r^{int}]
\end{aligned}
$$

Simplifying,

$$
\mathbb{E}[Q_r|X_r \in \mathcal{X}_r^{int}] = \frac{\bar{a}_t(r)\bar{q}_t(r) - \bar{a}_{t'}(r)\bar{q}_{t'}(r)}{\bar{a}_t(r) - \bar{a}_{t'}(r)} \tag{1.9}
$$

The intuition is presented graphically in Figure 1.1. The average outcome for the higher-intensity treator t is a weighted average of the conditional outcomes over the two signal groups. This identifies the average outcome conditional on an intermediate signal.

By (1.8), the conditional outcomes define intervals $S^{r,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}})$ of possible treatment costs $c_t(r)$ for each treator t and treatable group r. The test rejects the null that treator t is not prejudiced when:

$$
S^{M,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}}) \cap S^{W,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}}) = \emptyset.
$$

because this implies that $c_t(M) \neq c_t(W)$. Finally, the test can identify the victimized treatable group because the sets $S^{r,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}})$ (and therefore the costs of treatment) are strictly ordered on the real line if the test finds prejudice.

$\square$

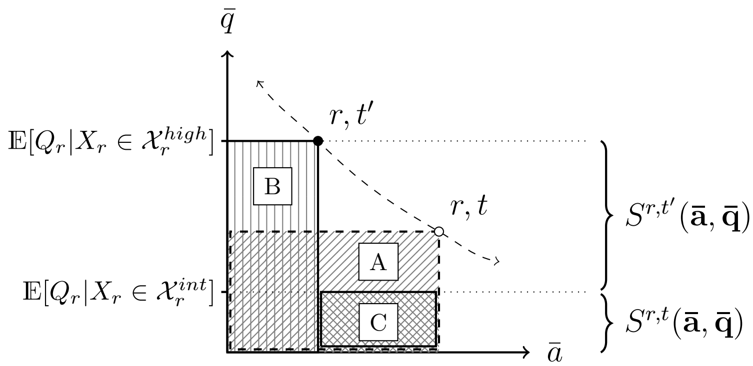

Figure 1.1. Fix a treatable group r. By the Law of Iterated Expectations, $E[Q_r|X_r \in \mathcal{X}_r^{int}]$ is identified as the solution to $A = B + C$. Since treator t' does not treat given any intermediate signal $X_r \in \mathcal{X}_r^{int}$, the average outcome given those signals provides a lower bound for $c_{t'}(r)$. Conversely, the average outcome over any set of signals a treator t does treat provides an upper bound for $c_t(r)$. The bounds define sets of possible search costs $S^{r,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}})$ for treator t on treated group r.

1.4.1. Relation to AF Test

Anwar and Fang (2006) propose a relative test for prejudice $T_R(\bar{\mathbf{a}})$ when individual returns to treatment are strictly diminishing:

$$\bar{a}_t(r) > \bar{a}_{t'}(r) \implies \bar{a}_t(r') > \bar{a}_{t'}(r'). \tag{1.10}$$

Average actions are perfect proxies for treatment costs in the AF model: higher average actions correspond to lower treatment costs for a fixed treatable group.[23] If (1.10) does not hold, it follows that treator t has lower treatment costs than t' over treatable group r but weakly higher treatment costs over treatable group r'. This implies that the treatment costs of at least one of the treators must depend on the treatable group – at least one of the treators is prejudiced.

[23] Average outcomes are also perfect proxies, so the test can be equivalently defined in terms of average outcomes $\bar{\mathbf{q}}$.

The power of a test is defined to be the probability of correctly rejecting the null hypothesis. The following theorem shows that the new test is more powerful than the test proposed by AF.

Theorem 1.2. *The absolute test T_A is more powerful than the AF relative test T_R.*

Proof: Suppose that returns to treatment are strictly diminishing, so that the AF test holds. It suffices to show that the test T_A rejects the null hypothesis of no prejudice if the test T_R does. Asymptotically, the test T_R finds prejudice if and only if the ordering of treatment costs across treators differs by treatable group. Therefore without loss of generality suppose that:

$$\begin{aligned} c_t(W) &\geq c_{t'}(W) \\ c_t(M) &< c_{t'}(M). \end{aligned}$$

For simplicity of notation, denote the bound used in the absolute test by:

$$\kappa_r \triangleq \mathbb{E}[Q_r | X_r \in \mathcal{X}_r^{int}], \quad r = M, W.$$

By the ranking of treatment costs and bounds from (1.7), it follows that:

$$\begin{aligned} c_t(W) \geq \;& \kappa_W \;\geq c_{t'}(W) \\ c_t(M) < \;& \kappa_M \;< c_{t'}(M). \end{aligned}$$

Note that the inequalities are strict in the second equation because returns to individual treatment are assumed to be *strictly* diminishing.

If $\kappa_M \leq \kappa_W$, then:

$$c_t(M) < \kappa_M \leq \kappa_W \leq c_t(W). \tag{1.11}$$

The absolute test rejects that treator t is unprejudiced because the intervals $S^{M,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}})$ and $S^{W,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}})$ are strictly ordered by (1.11). In the symmetric case where $k_M \geq k_W$, the absolute test finds prejudice for treators t'. $\square$

The new test T_A improves on the explanatory and statistical power of T_R. This increase in power comes at the cost of requiring more data. The AF test T_R requires data on actions *or* outcomes, while the new test T_A requires both. When the researcher observes both actions and outcomes (including in the AF dataset), however, the new test more efficiently uses the available information.

1.4.2. Relation to KPT Test

Knowles, Persico, and Todd (2001) propose a test $T_C(\bar{\mathbf{q}})$ for prejudice when returns to individual treatment are constant:

$$\bar{q}_t(M) = \bar{q}_t(W)$$

If marginal returns to individual treatment are constant everywhere, then average returns to treatment are also constant. The absolute test statistic reduces to the KPT test statistic to the extent that constant average returns are observed empirically. This makes sense because KPT *assume* that marginal – and therefore average – returns to individual treatment are constant everywhere.

Constant average returns to individual treatment are observed empirically when $\bar{a}_t(r) > \bar{a}_{t'}(r)$ and $\bar{q}_t(r) = \bar{q}_{t'}(r)$. Then the test inequalities (1.7) imply:

$$\bar{q}_{t'}(r) \geq c_{t'}(r) \geq \frac{\bar{a}_{t'}(r)\bar{q}_{t'}(r) - \bar{a}_t(r)\bar{q}_t(r)}{\bar{a}_{t'}(r) - \bar{a}_t(r)} = \bar{q}_{t'}(r)$$

$$\implies c_{t'}(r) = \bar{q}_{t'}(r).$$

In particular, for treator t', the new test statistic reduces to the average outcome – the same as the KPT test statistic.[24] In this sense KPT is a special case of the new test.

1.5. Introducing Differential Information

The previous analysis assumes that treators are *equally informed* over treatable groups.

Definition 1.5 ($t \sim_r t'$)**.** *Treators t and t' are* equally informed *over treatable group r if:*

$$F^{r,t} = F^{r,t'}. \tag{1.12}$$

However, this assumption is limiting. Treators may improve in their decision-making as they accumulate experience. In policing, officers may also be better at assessing guilt in motorists who are culturally or otherwise similar,[25] or motorists of a certain race may be more cooperative with certain groups of officers. Alternatively, as argued by Donohue and Levitt (2001), communities may be more willing to cooperate with an officer of the same

[24] Note that for treator t we can only infer that $c_t(r) \leq \bar{q}_t(r)$ because the data do not imply that marginal returns for treator t are constant.

[25] For example, a literature on cross-race facial recognition generally finds support for familiarity bias, or the notion that other-race groups "all look alike" [for early examples, see Malpass and Kravitz (1969); Brigham and Barkowitz (1978)].

race, who thus obtains more information about a motorist on the streets. As observed by AF, this makes the equal information assumption especially tenuous in policing if stops and searches occur on local streets.[26]

Differences in signal *informativeness* capture the previous examples. Therefore I consider an extension to the model where the informativeness of signals can vary across treators for a fixed treatable group. The natural notion of preferences over signal structures that satisfy FOSD is given by Athey and Levin (2017), Theorem 1, and replicated in Equation (A.1) in Appendix A.[27] If condition (A.1) holds I say that F' is strictly better than F and write $F' \succ F$. Say a signal F' is weakly better than F if it is strictly better or the same:

$$(F' \succsim F) = (F' \succ F) \vee (F' = F).$$

A treator is weakly better informed if the the treator receives a weakly better signal.

Definition 1.6 ($t \succsim_r t'$)**.** *Treator t is* weakly better informed *than treator t' over treated group r if:*

$$F^{r,t} \succsim F^{r,t'}. \tag{1.13}$$

[26] In health care, Bischoff et al. (2003) find evidence that language barriers between patients and nurses can impact symptom reporting and referrals.

[27] This condition is also implied by more demanding orderings of signal structures F, such as Blackwell (1951) and Lehmann (1988).

1.5.1. Test for Differential Information

The assumption that treators are equally informed is partially testable in the data. In this section I propose a new test for differential information across treators. This complements a testable implication provided by AF. The relation between the two conditions is presented graphically in Figure 1.2.

Theorem 1.3. *If treators t and t' are equally informed over treatable group r ($t \sim t'$), then:*

$$\bar{a}_t(r) > \bar{a}_{t'}(r) \implies \bar{a}_t(r)\bar{q}_t(r) \geq \bar{a}_{t'}(r)\bar{q}_{t'}(r). \tag{1.14}$$

Proof: Suppose that (1.14) is violated:

$$\Big(\bar{a}_t(r) > \bar{a}_{t'}(r)\Big) \wedge \Big(\bar{a}_t(r)\bar{q}_t(r) < \bar{a}_{t'}(r)\bar{q}_{t'}(r)\Big).$$

Then:

$$\mathbb{E}[Q_r|X_r \in \mathcal{X}_r^{int}] = \frac{\bar{a}_t(r)\bar{q}_t(r) - \bar{a}_{t'}(r)\bar{q}_{t'}(r)}{\bar{a}_t(r) - \bar{a}_{t'}(r)} < 0$$

which violates the fact that $\mathcal{Q} \subset \mathbb{R}_+$. $\square$

The new testable implication stems from the bounded rate of change in average outcomes: if two treators are equally informed, then a small difference in average actions implies a small difference in average outcomes. Consider the policing example. If an officer who searches 0.9% of drivers finds contraband 10% of the time, then an equally informed officer who searches 1% of drivers cannot find contraband only 5% of the time. At worst, the second officer searches an additional 0.1 p.p. of drivers who are never guilty.

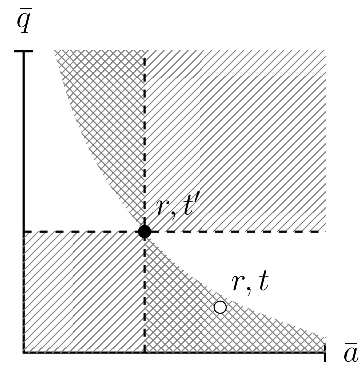

Figure 1.2. The lighter shaded region represents a rejection region of equal information derived by AF, and the darker area represents a new rejection region. In this case the new test would reject the assumption of equal information. The implied rate of change is too large in magnitude to be rationalized by the data if the treators are equally informed.

This places a lower bound of

$$\frac{0.9}{1}(10\%) + \frac{0.1}{1}(0\%) = 9\%$$

on the second officer's search success rate.

Condition (1.14) has an intuitive interpretation in the case of binary quality. In the policing example, (1.14) states that the *unconditional* probability of success[28] must be increasing in the search rate. Because a high-search officer searches a superset of the drivers searched by a lower-search officer, the overall probability of success for the high-search officer is at least as large as the overall probability of success for the low-search officer when the officers are equally informed.

[28] More precisely, *unconditional* on search but conditional on stop.

1.5.2. Test for Prejudice, Differential Informativeness

Suppose the researcher rejects the assumption that treators are equally informed. Differential information may be observed empirically or assumed by the nature of the application. For example, a more experienced police officer or doctor may be better informed than a less experienced counterpart. In many applications, the possible variation in information across treators is captured by considering differences in signal *informativeness*. This section presents a test for prejudice when signal informativeness varies across treators.

Theorem 1.4. *Suppose individual returns to treatment are weakly diminishing and that treator information over treated groups is given by a non-empty $\succsim$. There exists an* absolute *test $\tilde{T}_A(\bar{\mathbf{a}}, \bar{\mathbf{q}}; \succsim)$ for prejudice based on partial identification of treatment costs:*

$$c_t(r) \in \tilde{S}^{r,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}}; \succsim).$$

Specifically, if treator t is weakly better informed than treator t' over treatable group r, there exists a test $\tilde{T}^{r,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}}; \succsim)$ that treator t is prejudiced in favor *of treatable group r relative to treatable group r':*

$$\tilde{S}^{r,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}}; \succsim) > \tilde{S}^{r',t}(\bar{\mathbf{a}}, \bar{\mathbf{q}}; \succsim).\text{[29]} \tag{1.15}$$

Proof: Let $k^{r,t}(\bar{a})$ denote the possibility frontier of average outcomes that a treator t can achieve by taking average action $\bar{a}$ under signal structure $F^{r,t}$. Better information expands this possibility frontier:

[29] Where $S_1 > S_2$ iff $\inf S_1 > \sup S_2$.

Lemma 1.2. *Better signals expand the treator possibility frontier:*

$$F^{r,t} \succsim F^{r,t'} \implies k^{r,t}(\bar{a}) \geq k^{r,t'}(\bar{a}) \quad \textit{for all } \bar{a}.$$

The proof of Lemma 1.2 is provided in Appendix A.

The derivation of the absolute test under differential information proceeds analogously to the derivation under equal information, with the addition of the inequality from Lemma 1.2.

Suppose again that $\bar{a}_t(r) > \bar{a}_{t'}(r)$. Since treators are no longer equally informed, define:

$$\begin{aligned}
\mathcal{X}_{r,t}^{high} &= \{x | F_X^{r,t}(x) \geq 1 - \bar{a}_{t'}(r)\} \\
\mathcal{X}_{r,t}^{int} &= \{x | 1 - \bar{a}_t(r) \leq F_X^{r,t}(x) \leq 1 - \bar{a}_{t'}(r)\} \\
\mathcal{X}_{r,t'}^{high} &= \{x | F_X^{r,t'}(x) \geq 1 - \bar{a}_{t'}(r)\} \\
\mathcal{X}_{r,t'}^{int} &= \{x | 1 - \bar{a}_t(r) \leq F_X^{r,t'}(x) \leq 1 - \bar{a}_{t'}(r)\}
\end{aligned}$$

Similar bounds to those derived in (1.7) continue to hold for each officer:

$$\mathbb{E}[Q_r | X_{r,t'} \in \mathcal{X}_{r,t'}^{high}] \geq c_{t'}(r) \geq \mathbb{E}[Q_r | X_{r,t'} \in \mathcal{X}_{r,t'}^{int}] \tag{1.16}$$

$$\mathbb{E}[Q_r | X_{r,t} \in \mathcal{X}_{r,t}^{int}] \geq c_t(r). \tag{1.17}$$

The test loses partial power relative to the test under equal information at the next step of identifying the average conditional outcomes $\mathbb{E}[Q_r | X_{r,t} \in \mathcal{X}_{r,t}^{int}]$ in terms of observable average actions and outcomes.

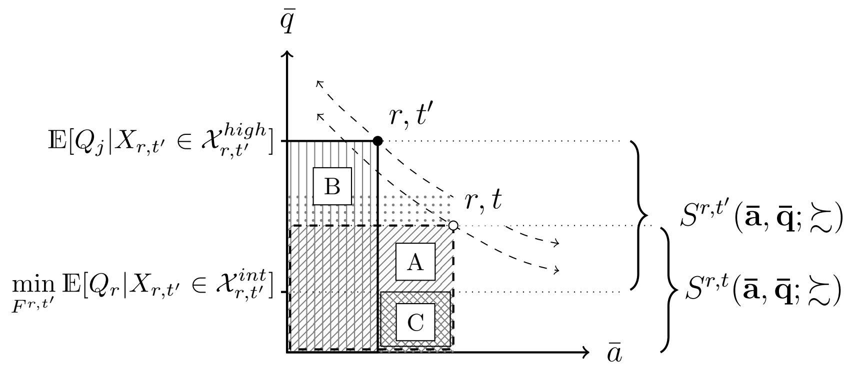

Figure 1.3. A more informative signal expands the outcome frontier and adds the dotted wedge to the reasoning under equal information. Therefore the Law of Iterated Expectations implies $A \leq B + C$, and only the minimum of $\mathbb{E}[Q_r|X_{r,j} \in \mathcal{X}^{int}_{r,j}]$ is identified for $j = t, t'$. This provides a bound for the better-informed treator. For the other treator, $S^{r,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}}; \succsim) = [0, \bar{q}_t(r)]$. In the case where the better-informed treator takes a higher average action, analogous reasoning identifies the maximum of $\mathbb{E}[Q_r|X_{r,j} \in \mathcal{X}^{int}_{r,j}]$ and therefore an upper bound.Note that in both cases the inter-marginal logic only identifies a bound for the better-informed treator.

If $F^{r,t} \precsim F^{r,t'}$, then:

$$\text{(By Lemma 1.2)} \quad \bar{q}_t(r) \leq k^{r,t'}(\bar{a}_t(r))$$

$$= \mathbb{E}\Big[Q_r|X_{r,t'} \in \mathcal{X}^{high}_{r,t'} \cup \mathcal{X}^{int}_{r,t'}\Big]$$

$$= \frac{\bar{a}_{t'}(r)}{\bar{a}_t(r)}\mathbb{E}\Big[Q_r|X_{r,t'} \in \mathcal{X}^{high}_{r,t'}\Big] + \frac{\bar{a}_t(r) - \bar{a}_{t'}(r)}{\bar{a}_t(r)}\mathbb{E}\Big[Q_r|X_{r,t'} \in \mathcal{X}^{int}_{r,t'}\Big]$$

$$= \frac{\bar{a}_{t'}(r)}{\bar{a}_t(r)}\bar{q}_{t'}(r) + \frac{\bar{a}_t(r) - \bar{a}_{t'}(r)}{\bar{a}_t(r)}\mathbb{E}\Big[Q_r|X_{r,t'} \in \mathcal{X}^{int}_{r,t'}\Big] \tag{1.18}$$

A graphical version of the argument is presented in Figure 1.3. Differential signal informativeness introduces a wedge into the argument under equal information. Even though observables now only partially identify the average inter-marginal outcome, the bound on the treatment cost for the better-informed treator remains unchanged as a function of observables. Combining (1.16) and (1.18),

$$c_{t'}(r) \geq \mathbb{E}[Q_r | X_{r,t'} \in \mathcal{X}^{int}_{r,t'}] \geq \frac{\bar{a}_t(r)\bar{q}_t(r) - \bar{a}_{t'}(r)\bar{q}_{t'}(r)}{\bar{a}_t(r) - \bar{a}_{t'}(r)} \tag{1.19}$$

For the less-informed treator,

$$\begin{aligned}
\bar{q}_t(r) &= \mathbb{E}\Big[Q_r \Big| X_{r,t} \in \mathcal{X}^{high}_{r,t} \cup \mathcal{X}^{int}_{r,t}\Big] \\
&= \frac{\bar{a}_{t'}(r)}{\bar{a}_t(r)} \mathbb{E}\Big[Q_r \Big| X_{r,t} \in \mathcal{X}^{high}_{r,t}\Big] + \frac{\bar{a}_t(r) - \bar{a}_{t'}(r)}{\bar{a}_t(r)} \mathbb{E}\Big[Q_r \Big| X_{r,t} \in \mathcal{X}^{int}_{r,t}\Big] \\
&= \frac{\bar{a}_{t'}(r)}{\bar{a}_t(r)} k^{r,t}(\bar{a}_{t'}(r)) + \frac{\bar{a}_t(r) - \bar{a}_{t'}(r)}{\bar{a}_t(r)} \mathbb{E}\Big[Q_r \Big| X_{r,t} \in \mathcal{X}^{int}_{r,t}\Big] \\
\text{(By Lemma 1.2)} \qquad &\leq \frac{\bar{a}_{t'}(r)}{\bar{a}_t(r)} \bar{q}_{t'}(r) + \frac{\bar{a}_t(r) - \bar{a}_{t'}(r)}{\bar{a}_t(r)} \mathbb{E}\Big[Q_r \Big| X_{r,t} \in \mathcal{X}^{int}_{r,t}\Big].
\end{aligned}$$

While this argument also provides a lower bound on the inter-marginal outcome $\mathbb{E}[Q_r | X_{r,t} \in \mathcal{X}^{int}_{r,t}]$ in terms of observables, the inequality is in the wrong direction to provide a meaningful bound on the treatment cost $c_t(r)$. Under the reverse assumption $F^{r,t} \succsim F^{r,t'}$ on signal informativeness, the inequalities are flipped. Therefore the argument provides an upper bound for $c_t(r)$ but not a lower bound for $c_{t'}(r)$.

Thus the bounds from the equal information case are maintained *only* for a weakly better-informed treator:

$$t \succsim_r t' \implies S^{r,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}}) = \tilde{S}^{r,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}}; \succsim)$$

In the absence of better information,

$$S^{r,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}}) \subseteq \tilde{S}^{r,t}(\bar{\mathbf{a}}, \bar{\mathbf{q}}; \succsim) = [0, \bar{q}_t(r)]$$

Weakening the equal information assumption to a ranked informativeness assumption leads to a weaker absolute test due to weaker bounds on treatment costs. Note that the researcher can never reject a null hypothesis of no prejudice for a treator with only upper bounds on treatment costs – any sufficiently low treatment cost rationalizes the observed behavior without reference to prejudice. Therefore a researcher can never reject a null hypothesis of no prejudice for a treator who is not assumed to be at least weakly-better informed over some treated group. In particular, a treator must be assumed weakly-better informed over a treatable group r in order to test whether the treator is prejudiced *in favor* of treatable group r. $\square$

1.6. Applications

1.6.1. Motor Vehicle Searches: Anwar and Fang

In this section I revisit the empirical application of Anwar and Fang (2006). The AF dataset provides an ideal opportunity to apply the new absolute test because the model underlying the new test encompasses the AF model. This confers two advantages. First, the dataset suggests that marginal returns to treatment intensity are decreasing, which is a defining feature of the model. Second, because the model of officer behavior is held fixed across the tests T_A and T_R, any difference in findings stems from the efficiency of the new test, rather than stronger assumptions.

The dataset consists of 906,339 stops and 8,976 searches conducted by the Florida State Highway Patrol from January 2000 to November 2001. For convenience, Tables 1.1 and 1.2 replicate the search and search success rates by officer and motorist race presented in Table 1 of their paper. Note that this is also the set of summary statistics used by AF in applying their tests.[30]

Applying the Absolute Test T_A. To apply the absolute test, bounds are computed for each pair of officers who are adjacent in the space of search and search success rates. The computed bounds are presented numerically and graphically in Figure 1.4. Note that the bounds are not confidence intervals, but rather partial identification of the parameter of interest, the vector of search costs.[31] Note that the new test provides partial bounds on search costs for *each* group of officers.

In practice, the absolute test finds evidence that Hispanic officers are prejudiced against Hispanic motorists relative to black and white motorists. This is the first formal evidence of prejudice in the AF dataset. Thus the new test T_A also confers additional explanatory and statistical power relative to T_R *in practice.* Even when the test fails to find prejudice, partial identification of parameters can provide useful bounds on the possible extent of prejudice. For example, given the sharpness of the bounds, the absolute test suggests that any prejudice exhibited by Hispanic officers between black and white motorists is limited.

Conclusions are sharpest for Hispanic officers in part because of their intermediate rank in search intensity. Because they search each motorist group at a higher rate than black officers, the test inequalities provide new upper bounds on search costs. Because they search less than white officers, the inequalities also provide lower bounds. The other two

[30] In particular, AF apply a re-sampling procedure in order to ensure that officers of different races are stopping the same pool of motorists in expectation. More detail on the procedure is provided in Section III, C of their paper.
[31] Confidence intervals on the bounds can also be computed. For methods in the case of partially identified parameters, see Manski (2003) or Manski and Imbens (2004).

officer groups have either sharp upper or lower bounds. In particular, an absence of lower bounds precludes any finding of prejudice for white officers because the data can always be rationalized by assuming that white officers have a sufficiently low search cost across all motorist groups. Insofar as the researcher is most interested in unearthing prejudice among the officers with the highest propensity for search, this is a general limitation of the test. On the other hand, the concern is unique to absolute tests, since a relative test can never identify prejudice in a particular group by definition. Additional variation across officers, such as gender or experience effects, can ameliorate this identification issue, as well as provide more disaggregated tests for prejudice.

Table 1.1. [AF] Search Rate Given Stop (%)

	Trooper Race		
Driver Race	White	Black	Hispanic
White	0.96	0.27	0.76
Black	1.74	0.35	1.21
Hispanic	1.61	0.28	0.99

Table 1.2. [AF] Search Success Rate (%)

	Trooper Race		
Driver Race	White	Black	Hispanic
White	24.3	39.4	26.0
Black	19.9	26.0	20.8
Hispanic	8.5	21.0	14.3

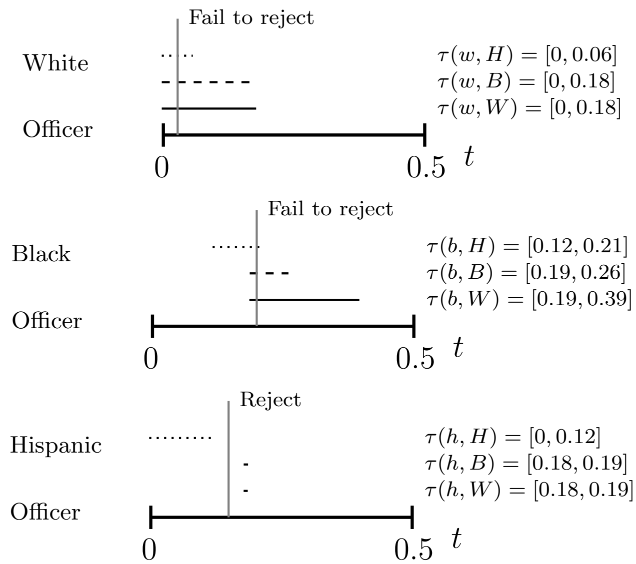

Figure 1.4. The bounds provide evidence of prejudice by Hispanic officers against Hispanic motorists because the sets of possible search thresholds do not overlap. On the other hand, a hypothesis of no prejudice can be rationalized with a search threshold that intersects the sets across motorist groups. Therefore the test fails to reject the null hypothesis of no prejudice by white and black motorists.

Applying the Test for Unequal Information. The new information test also provides the first evidence in the AF data that officers are not equally informed. Recall that under equal information, the unconditional probability of success should be increasing in the search rate. However, white officers search 1.61% of stopped Hispanic motorists and find contraband 8.5% of the time that they conduct a search. Hispanic officers search 0.99% of stopped Hispanic motorists and find contraband 14.3% of the time that they conduct a search. Therefore Hispanic officers find more contraband conditional on stopping Hispanic motorists despite searching less:

$$0.99 \times 14.3 = 14.157 > 13.685 = 1.61 \times 8.5$$

This implies that Hispanic and white officers are not equally informed over Hispanic motorists.

The rejection of equal information suggests that Hispanic officers may receive better signals of guilt over Hispanic motorists than white officers. Note that as long as this finding does not alter the researcher's stance on informational assumptions for the other groups, the finding of heterogeneity does not affect the outcome of the test presented in the previous section. In particular, the finding of prejudice holds as long as the researcher assumes that Hispanic officer signals are at least as informative as white officer signals for each motorist group.[32]

1.7. Conclusion

Racial and ethnic disparities remain prevalent across a range of domains. Distinguishing among the possible sources of such disparities matters. This underscores the value of new methods for determining whether disparities are driven by statistical discrimination or prejudice. The paper develops the first absolute test for identifying prejudice in the setting where treators make binary treatment decisions and where the returns to extending treatment to a larger fraction of the population are weakly diminishing. For example, employers hire, judges deny bail, law enforcement officers engage in stops and searches, doctors administer procedures and medical tests, and creditors extend loans and mortgages.

The model unifies the existing literature by encompassing the environments of existing models. The test provides a more powerful alternative to the test derived in AF when

[32] On the other hand, assuming that Hispanic officer signals are *not* at least as good over, say, white motorists would affect the finding of prejudice against Hispanics only relative to white motorists.

marginal returns to treatment are diminishing. This allows the test to uncover the first evidence of prejudice in the AF dataset. The test also includes the KPT test statistic as a special case when the empirical evidence supports the assumption of constant marginal returns to individual treatment intensity. Methodologically, the test is the first to jointly use actions *and* outcomes in the test statistic. A weakened version of the new test holds even when treators may vary in the informativeness of their signals. Such improvements come at the cost of more stringent but frequently non-binding data requirements.

CHAPTER 2

Revenue from Matching Platforms (joint with James Schummer)

2.1. Introduction

The proliferation of online platforms has led to intensified interest in the study of platform pricing. The topic becomes increasingly important—particularly for regulators—as dominant platforms emerge in "winner-take-all" environments.[1] While an existing literature tells us much about pricing on certain kinds of platforms, much of it has ignored two market characteristics that, together, distinguish some platforms from those that have been studied. Specifically, we consider platforms serving markets in which *exclusive* (one-to-one) partnerships occur between *horizontally differentiated* agents.

Exclusivity. Canonical models of platforms (see Subsection 2.1.2) address *many-to-many* matching environments, where each participating agent interacts with *all* the agents on the other side of the market. These models accurately portray often-cited examples of platforms such as credit cards (connecting consumers and merchants), video game consoles (game players and developers), and newspapers (readers and advertisers). On the other hand, many platforms exist specifically to create *one-to-one* matchings, such as AirBnB (guests lodging with hosts), Uber (riders and drivers), and online dating platforms. Each AirBnB guest wants to be matched to a single host, and each host desires a single guest;

[1] E.g. The Economist, "Nostrums for rostrums," May 26, 2016.

Uber drivers and passengers also are matched one-to-one at any given point in time. [2] For people interested in developing a monogamous relationship, the outcome ideally produced by a dating platform is a one-to-one matching.[3]

Heterogeneity. There remains a gap between the literature on platform pricing and the literature pioneered by Gale and Shapley (1962) concerning capacity-constrained matching of agents with heterogeneous tastes. For example, canonical papers in the former literature address agents who would obtain the same value from *all* partners to whom they are matched; agents on one side of the market perceives the other side's agents as homogeneous objects. Again, this assumption makes sense in certain applications—credit card holders and merchants value a credit card based on its "cashless" feature rather than the identity of their matched partners. In one-to-one markets, however, this homogeneity is atypical. Instead, an agent on any one side (e.g. an AirBnB guest) typically perceives the other side's agents (AirBnB hosts) as heterogeneous objects. Furthermore, these heterogeneous tastes may differ even across agents within one side of the market, creating horizontal differentiation amongst the agents on the other side. AirBnB guests have different tastes over the type and location of residences; hosts have different preferences over guests with pets, children, or other needs.[4]

[2] Slightly complicating our story, Uber also offers a pooled service in which customers share a driver. There could be analogous exceptions to the one-to-one rule on AirBnB. Nevertheless, the capacity constraints implicit on these platforms are better approximated with one-to-one models than the "all-to-all" models in the canonical literature.

[3] This is true even if agents have to "learn their preferences" in the short run by initially dating multiple people through the platform. If a single, long term relationship is the ultimate goal then a one-to-one model captures the essence of the platform.

[4] Numerous articles on the internet advise AirBnB hosts on how to accept or reject guest applications. The subjectiveness of these articles merely strengthens our opinion that tastes vary.

To consider the pricing implications of these two characteristics, we consider the revenue maximization problem for a monopolistic platform that sets prices to two sides of a "marriage market" and creates a one-to-one matching between the two sides. As a minimal requirement, we restrict attention to platforms that are (ex post) *individually rational*: once agents anticipate the outcome of the platform's matching process, they should not want to renege on participation and payment. Given our interest in heterogeneous preferences, we also consider the revenue and pricing consequences for a platform that further commits to producing *stable* outcomes, i.e. ruling out the ex post possibility that a "blocking pair" of agents would have preferred being matched with each other, *even when taking the platform's pricing into account.* Our interest in the stability condition is motivated both normatively and positively.

Normatively, a platform may wish to use stable matching procedures for a variety of reasons. For one, stability allows a matchmaker to advertise the "quality" of its matching procedure in that no (blocking) pair of customers could come to the realization that they could have jointly created a better match than the one created by the platform.[5] A second reason is that stability has been argued to help prevent market unravelling; see Roth (2002) for evidence supporting this argument. Even if a monopolistic platform has technology that, in the short run, gives it exclusive control over agents' ability to match, its creation of *unstable* matchings could increase the platform's long run vulnerability to the entry of alternate technologies/platforms that would allow blocking pairs to match outside

[5] For example, online matchmaker eHarmony claims to use its patented algorithms to "identify matches with the highest potential for a successful relationship" and to predict "who you match best with." Since these promises are being made to *both* sides of the market, one can interpret these claims as an attempt to determine stable outcomes.

of the monopolist's current platform. A farsighted, monopolistic matching platform could thus consider stability to be a form of entry deterrence.

To see the positive motivation for studying stability in this problem, consider platforms (e.g. dating sites) that "create" matchings in a decentralized way, by allowing the agents themselves to form pairs. In such settings, stable outcomes can arise naturally [Hitsch et al. (2010)]. By allowing decentralized matching, the platform has indirectly committed to providing a stable outcome. In such markets, stability should be viewed more as a market design constraint than as a choice.

2.1.1. Overview of results

Our focus is on how a platform would set prices to two sides of a matching market as a function of market parameters. These parameters include not only the distribution of individual agent's values, but also the degree of market imbalance (relative sizes of the two sides) and the degree of correlation in agents' values. While imbalance and correlation play little role in standard models of many-to-many platforms, these characteristics *can* impact pricing in our setting (see Section 2.5).

Our first set of results concerning market imbalance appear counterintuitive with respect to a recent result in the literature on stable matching. In a classic marriage model with independent, uniformly drawn preferences, Ashlagi et al. (2017) show the following striking result. Under essentially any degree of market imbalance, average (normalized) payoffs are notably higher for agents on the short side of the market than for those on the long side, at *any* stable outcome. Consider the following intuition. A woman walks into a balanced market. If the existing stable match is "far" from the male-optimal stable match in terms of payoffs (for example, the woman-optimal stable match), the woman easily breaks a match. The recently unmatched woman is in turn also likely to break a match. This pattern of match-breaking propagates until men begin to derive high value from their current match. Otherwise the probability that the lone woman will break a match remains high. Ashlagi et al. (2017) show formally that the payoffs under the match described above are close to the payoffs under the man-optimal match. Therefore the payoffs under *any* stable match are close to those under the man-optimal match.

Following this result one might guess that a monopolist who *controls* the stable matching process should use unbalanced prices in order to capture these unbalanced payoffs, charging a higher price to the short side of the market. Under such fees, we show that this reasoning does not hold: We prove a "symmetric-pricing" result stating that a revenue maximizing stable platform *does not* price discriminate between the two sides of the market based on their relative sizes, despite the asymmetric-payoff result of Ashlagi et al. Particularly, when agents' values are i.i.d. across *both* sides of the market a standard hazard rate condition leads the platform to charge the same price to both sides of the market, regardless of their relative sizes.

At first glance our no-price-discrimination result may appear to coincide with results obtained for the many-to-many platform models described in Subsection 2.1.2. In those models it is conventional wisdom to subsidize the price-sensitive side of the market in order to exploit cross-network effects.[6] Therefore it may sound unsurprising when we show an absence of price discrimination in our model when both sides are equally price sensitive. This is a misleading comparison, however. In many-to-many models, this argument not to

[6] A small price drop on the sensitive side leads to a large increase in transactions, proportionally increasing revenue earned from the other side.

price discriminate applies *regardless* of whether agent's values are correlated. Intuitively, in many-to-many models a platform is essentially pricing each transaction separately, so correlation plays no role in maximizing expected revenue across the entire market. On the other hand, when values become correlated in capacity constrained models such as ours, market imbalance *can* affect the platform's optimal pricing decision, leading to different pricing across the two sides.[7] These ideas and comparisons are explained in Section 2.5.

The combination of these results illustrates that our model sits between the literatures on matching and on platform pricing. Despite the fact that, with i.i.d. values, the short side of the market captures more value than the long side under stability [Ashlagi et al. (2017)] a monopolistic platform does not capture value by charging a premium to the short side. At the same time, if market imbalance coexists with *correlation* in agents' values, the platform does price discriminate on the basis of market imbalance, a phenomenon that would not occur in many-to-many models without capacity constraints.

Our tool for analyzing stable platform pricing is a new family of (typically unstable) matching procedures that we call *Meet and Propose* (MAP) algorithms. These algorithms work similarly to *Deferred Acceptance* algorithms, but require proposers to "meet" potential mates in some arbitrarily fixed order. We use the relationship between MAP algorithms and Deferred Acceptance to prove the no-price-discrimination theorem described earlier.

As part of our analysis of MAP algorithms, we derive a closed-form expression that approximates a stable platform's expected revenue, when agents have i.i.d. values. This

[7] However the *direction* in which the platform price discriminates turns out to depend on the type of correlation in values, so again the reasons for price discrimination are not related to the surplus imbalance results of Ashlagi et al.

is accomplished by considering the special case in which MAP algorithms coincide with a "constrained-dictatorship" mechanism, representing a platform where agents from one side arrive sequentially and are matched with their favorite (mutually compatible) remaining agent on the other side.[8] By appealing both to simulations and a related result of Arnosti (2016), we argue that our approximation is a fairly tight lower bound on revenue when a platform uses stable matching mechanisms in large markets.

We also use this bound to show that, as markets with independently drawn preferences grow arbitrarily large, the platform's "cost of committing to stability" essentially vanishes: that is, by relaxing stability to merely ex-post individual rationality, a platform would improve its revenue by only a vanishingly small percentage. Hence our lower bound on the stable platform's revenue turns out to be a good approximation of that revenue as the market becomes large. Furthermore, simulations demonstrate that the bound is a good approximation of revenue for markets of *any* size.

To conclude, we consider the implications of correlation in agents' values. Correlation in the one-to-one matching case introduces incentives for the platform to price-discriminate based on relative market size. Furthermore, the type of correlation matters. Two-sided matching suggests two dimensions of correlation, which have opposite effects on relative pricing comparative statics. On the other hand, correlation does not affect pricing decisions in the many-to-many case. This suggests that the impact of preference correlation on pricing decisions varies with the amount of intra-platform competition.

[8] Though it is not our objective, our analysis can be applied directly to such platforms, including those where pairwise compatibility is determined exogenously.

2.1.2. Platforms Literature

The two-sided markets literature shares our interest in profit-maximizing platforms that connect two sets of agents who derive benefits from interacting with each another. In contrast to our approach many models in this literature exhibit the following features, making them less relevant for the applications we have in mind.

All-to-all:: each agent receives some constant marginal benefit from each agent present on the other side of the market;

No differentiation:: each agent perceives the agents on the other side of the market as indistinguishable;

No same-side externalities:: each agent is unaffected by the presence of other agents on the same side.

These features are implied by the modeling assumption that an agent's payoff is some affine function, say $a + bn$, of the number of agents on the other side of the market, n. The fixed and per-transaction benefits, a and b, may or may not be assumed to vary across agents.[9]

The affine payoff structure easily captures so-called "cross-network effects" where agents benefit from additional participation on the other side. A consequence is one of the fundamental lessons from the two-sided markets literature: a profit-maximizing platform should not set prices to the two sides of the markets independently, as if they were two unrelated products. An increase in the per-transaction price charged to one side of the market can be viewed as a decrease in the marginal cost of providing transactions

[9] This payoff structure is used by the seminal papers of the literature, such as Rochet and Tirole (2003),(2006) and Armstrong (2006), as well as Caillaud and Jullien (2003), Weyl (2010), and others.

to the other side of the market [Rochet and Tirole (2003)]. This observation yields the *see-saw effect* observed by Rochet and Tirole (2006): if the platform finds reason to decrease the price charged to one side of the market (e.g. due to increased price sensitivity on that side), then this in turn justifies a price *increase* to the other side of the market.

A form of cross-network effect also exists in the one-to-one matching markets we consider. A well-known result of Gale and Sotomayor (1985a) states that the addition of an agent on one side of the market weakly improves stable outcomes for agents on the other side of the market. Since agents are heterogeneous and are capacity constrained to match with at most one other agent, however, the derivation of an analogous see-saw effect in this model is not as transparent, though its intuition is essentially the same.

On the other hand, one-to-one matching models distinguish themselves from two-sided market models in that the latter typically ignore *market size effects*.[10] In fact, the *all-to-all* feature of these latter models is what neutralizes market-size effects outright. In our one-to-one model, on the other hand, market size *does* impact pricing; furthermore its effect is determined by the degree of correlation in agents' preferences. See Section 2.5 for elaboration.

Another feature of the literature—beyond our current interest—is the analysis of competing platforms. The "divide and conquer" theme arising from this literature formalizes the idea that platforms may subsidize a "critical" side of the market and recover profits from the other. As Armstrong (2006) points out, if agents on only one side of the market must "single-home" (commit to a single platform), then platforms will compete for

[10] An exception is the distantly related work of Ambrus and Argenziano (2009). They consider competing platforms where one is cheaper and larger on one side of the market, and the other is cheaper and larger on the other side. This (equilibrium) price discrimination causes "high value" agents to choose expensive platforms in order to access more potential partners.

them (through low/subisidized pricing) while charging the “multi-homing” side monopoly prices.

Somewhat closer to our work, Damiano and Li (2008) consider competing platforms where heterogeneous agents are randomly matched one-to-one and each payoff is the product of the pair’s types.[11] One of their main points is that, in their setup, prices lead to an assortative segmentation of agents across platforms. Besides the issue of platform competition, there are other critical differences between their model and ours. Whereas our main results concern the case of heterogeneous preferences, Damiano and Li’s agents have the same (ordinal) preferences over the “vertically differentiated” agents on the other side; however we do consider perfectly correlated (cardinal) preferences in Section 2.5. Second, their random matching assumption sets aside market size effects, which is natural since their motivation is the question of market composition or quality; on the other hand we obtain such effects precisely when there is correlation in preferences. Finally, as in the literature described above, there are no same-side externalities: holding all else constant, an agent does not suffer from the presence of additional agents on the same side. This is not the case in our model [Gale and Sotomayor (1985a)].

Fershtman and Pavan (2016) study many-to-many matching platforms with persistent, private information. Their analysis requires a dynamic model, since agents remain in the market forever, potentially changing partners as their match values evolve. Optimal mechanisms are shown to be dynamic auctions that assign matches based on a scoring rule that uses reported information about both the persistent and idiosyncratic part of agents’ types.

[11] Also see Damiano and Li (2007) for a similar monopolistic setting. Results here deal with efficiency.

2.2. Model

The agents consist of a set $\mathcal{M}$ of men and a set $\mathcal{W}$ of women. A (one-to-one) **matching** is a function $\mu\colon \mathcal{M} \times \mathcal{W} \to \mathcal{M} \times \mathcal{W}$ satisfying the following usual conditions for all $m \in \mathcal{M}, w \in \mathcal{W}$: $\mu(m) \in \mathcal{W} \cup \{m\}$, $\mu(w) \in \mathcal{M} \cup \{w\}$, and $\mu(m) = w$ if and only if $\mu(w) = m$. We say that agent $i \in \mathcal{M} \times \mathcal{W}$ is unmatched (or single) when $\mu(i) = i$.

If man $m \in \mathcal{M}$ is matched to woman $w \in \mathcal{W}$, he obtains **value** $u_m(w) \in [0,1]$; similarly w obtains $u_w(m) \in [0,1]$ from being matched to m. The value of being unmatched is normalized to zero (denoted $u_i(i) \equiv 0$ when necessary). Each value $u_m(w)$ is randomly drawn according to a distribution F_M, and each $u_w(m)$ is drawn according to F_W, where both densities are continuously differentiable with positive support on $[0,1]$. We initially assume that each value $u_i(j)$ is drawn independently of all other values. Correlated values are considered in Section 2.5.

Though we rule out transfers between agents, they may make payments to the platform itself. Agents' preferences are represented by payoffs that are quasi-linear in such payments. Specifically, at a matching μ, if man $m \in \mathcal{M}$ makes a payment of p_M to the platform, then his payoff is $u_m(\mu(m)) - p_M$. The symmetric assumption holds for women.

We assume that the platform can price discriminate based only on (i) whether the agent is matched or single, and (ii) the agent's side of the market ($\mathcal{M}$ vs. $\mathcal{W}$). Fixing $\mathcal{M}$ and $\mathcal{W}$, **prices** are simply a pair of "match-contingent fees" $(p_M, p_W) \in \mathbb{R}^2$, where matched men and women are charged p_M and p_W respectively, while the payments of unmatched agents are normalized to zero. Implicitly, prices can be a function of (the size of) the sets $\mathcal{M}$ and $\mathcal{W}$, but we do not require such notation.

2.2.1. Constraints of the Platform

Our main interest is in platforms that charge agents fees for being part of a *stable* match [Gale and Shapley (1962)], i.e. a match guaranteeing individual rationality and the absence of pairwise blocking. Stability can be viewed either as a normative constraint or as a positive description of platform outcomes. Normatively, a platform designer might wish to avoid creating a blocking pair simply to avoid attempts by agents to manipulate or bypass the platform or to avoid complaints from dissatisfied customers. Positively, even if the matchmaker has little control over *how* agents match after they join the platform, stable matchings might arise naturally. For example Hitsch, Hortaçsu, and Ariely (2010) show that matchings in decentralized, online dating markets are approximately stable even though such platforms clearly cannot dictate which matchings occur.

Observe that the stability condition becomes endogenous once the platform can vary prices. On the other hand one can think of any *fixed* price as simply representing an agent's outside option (from remaining single), in which case the standard definition of stability applies. Regardless, we now formalize the concept as functions of prices $p = (p_M, p_W)$.

The individual rationality requirement means that an agent should not prefer to withdraw from the platform. Here, this means that no matched agent should prefer remaining single (for free) over his/her match outcome *given the platform's prices.*[12]

Definition 2.1. *Fix values u. With respect to prices $p = (p_M, p_W)$, a matching μ is* ***individually rational at p*** *when, for all $m \in \mathcal{M}$ and $w \in \mathcal{W}$, $\mu(m) = w$ implies both $u_m(w) \geq p_M$ and $u_w(m) \geq p_W$.*

[12] This is an *ex post* individual rationality condition, which is stronger than a definition which would only require agents to benefit from the platform in expectation, *ex ante*. There turns out to be little difference between these definitions in our model.

The no-blocking-pairs requirement means that no man-woman pair would strictly prefer matching with each other instead of receiving their prescribed matching outcome. In our context we are ruling out blocking pairs would would benefit from being matched with each other, *even if they have to pay the platform's prices in order to do so.*

Definition 2.2. *Fix values u. With respect to prices $p = (p_M, p_W)$, man $m \in \mathcal{M}$ and woman $w \in \mathcal{W}$* ***p-block*** *a matching μ when*

$$u_m(w) - p_M > u_m(\mu(m)) - p_M * \mathbb{1}_{\mu(m)\in W} \text{ and}$$

$$u_w(m) - p_W > u_w(\mu(w)) - p_W * \mathbb{1}_{\mu(w)\in M}$$

where $\mathbb{1}$ is the indicator function.

Combining these definitions leads to the following notion of stability.

Definition 2.3. *Fix values u. With respect to prices $p = (p_M, p_W)$, a matching μ is* ***p-stable*** *if (i) it is* individually rational at p *and (ii) there is no m, w who p-block μ.*

It is immediately clear that a p-stable matching can be found by simply (i) truncating each man's (woman's) preferences by removing all potential mates who are valued at less than p_M (p_W), then (ii) running the classic Deferred Acceptance (DA) algorithm on these truncated preferences. Furthermore it follows from well-known results [Roth (1984b)] that all p-stable matchings contain the same number of marriages.[13]

Theorem (Rural Hospital Theorem)**.** *Fix values u and prices $p = (p_M, p_W)$. All p-stable matchings contain the same number of marriages.*

[13] The result furthermore states that the set of married agents remains constant across all p-stable matchings. See also McVitie and Wilson (1971).

2.2.2. Implicit Informational Assumptions

In our analysis we take each realization of agents' preferences u as given and then compute a realized matching as a function of these preferences. When preferences can be observable by the platform, then this is without loss of generality. However we can justify this assumption even for environments in which preferences are not observable to the platform using results from the literature.

For instance, if the platform's objective is to yield p-stable matchings, then under some assumptions this can be done by setting up the proper "game." This is certainly the case when agents have complete information about each others' preferences. Specifically, Roth (1984a) shows that under the Deferred Acceptance revelation game, a (complete information) equilibrium outcome in undominated strategies must be stable. Kara and Sönmez (1996) show that the set of stable outcomes is fully implementable in Nash equilibrium. An even sharper prediction of stability is made using iterated deletion of dominated strategies; see Alcalde (1996).

Under some of these results we cannot necessarily predict *which* stable outcome would be obtained. In our setting this is irrelevant of course since the Rural Hospital Theorem implies that the platform does not care—the same number of marriages are created at any stable matching, so all such matchings provide identical revenue to the platform.

A second argument in favor of our informational assumption comes from a growing body of work confirming the idea that, in large markets, agents have little incentive to misreport their preferences [Roth and Peranson (1999); Immorlica and Mahdian (2005); Kojima and Pathak (2009); Azevedo and Budish (2015)].

Finally, as mentioned earlier, there is empirical evidence supporting the idea that in certain decentralized markets, stable outcomes obtain as real world "equilibrium" outcomes; e.g. Hitsch et al. (2010). For all of these reasons, we make the simplifying assumption that a p-stable matching is obtained by the platform once agents' preferences are randomly determined.

2.3. Meet and Propose

As mentioned earlier, a p-stable matching can be found by running the Deferred Acceptance (DA) algorithm on agent's preferences that have been appropriately truncated with respect to prices p. Here we provide an alternative formulation of such a method, DA_p, which we describe in a somewhat atypical fashion. The reason for this formulation is to make clear how a second algorithm, used below in our analysis, relates to this one. To state it, for any prices $p = (p_M, p_W)$ we say that $m \in M$ and $w \in W$ are **p-compatible** if $u_m(w) > p_M$ and $u_w(m) > p_W$.

Definition 2.4 (DA_p algorithm)**.** *The algorithm takes values u as input and initializes all men to be single. The following two steps are executed in rounds $t = 1, 2, \ldots$ until each man is either matched or has "met" every woman.*

- **Step t.1::** *Each man m who is single "meets" his favorite woman w among those he has not already proposed to. (If no such women exist, he remains single.) He proposes to her if and only if they are p-compatible.*
- **Step t.2::** *Each woman becomes matched to her favorite man among those who have proposed to her. (If none exist, she remains single.) All other men become (or remain) single.*

It should be clear that DA_p is the standard DA algorithm, with a bit of redundancy in the notion of men first "meeting" women before proposing. In the typical description of DA, one can dispense with this since men only "meet" the women with whom they might be compatible in the first place. Any further incompatibility on the woman's part is taken care of when she decides whether to accept his offer. We have included this extra step in our description in order to highlight the difference between DA_p and the class of $\mathrm{MAP}_p^{\rhd}$ algorithms that we use to approximate the platform's revenue under p-stability.

Specifically, we consider algorithms where a man's "meeting order" is not perfectly aligned with his preferences as it is under Steps $t.1$ of DA_p. This separates meeting orders from preference orders, but proposals themselves are still tied to preferences through each man's "decision" of whether to propose.[14] Women, on the other hand, accept and reject proposals as they do under DA: they reject all proposals other than the best one they have received so far.[15]

To describe our algorithms, a **meeting order** for man $m \in \mathcal{M}$, $\rhd_m$, is a linear order on $\mathcal{W}$. A profile of meeting orders is denoted $\rhd = (\rhd_m)_{m \in M}$.

Definition 2.5 ($\mathrm{MAP}_p^{\rhd}$ algorithm)**.** *The algorithm is parameterized by a meeting order profile* $\rhd$*. It takes values* u *as input and initializes all men to be single. The following two steps are executed in rounds* $t = 1, 2, \ldots$ *until each man is either matched or has "met" every woman.*

[14] The word *decision* should not be taken literally: we are merely using this algorithm as a tool. It does not share the incentive properties of DA.

[15] Observe that since only p-compatible men make proposals, we need not check women's individual rationality constraints.

Step t.1:: *Each man m who is single "meets" the woman w ranked highest under $\rhd_m$ among those he has not already proposed to. (If no such women exist, he remains single.) He proposes to her if and only if they are p-compatible.*

Step t.2:: *Each woman becomes matched to her favorite man among those who have proposed to her. (If none exist, she remains single.) All other men become (or remain) single.*

Fixing values u, if each $\rhd_m$ happens to coincide with the ordering of m's relative preference order over women ($u_m()$), then $\mathrm{MAP}_p^{\rhd}$ would yield the same outcome as DA_p for u. Typically, of course, a randomly determined u induces orders over women that do not coincide with $\rhd$, and the algorithms yield different outcomes. Nevertheless, we use $\mathrm{MAP}_p^{\rhd}$ to analyze and approximate the *distribution* of p-stable marriages in a random economy.

Finally, observe that we have described "all men propose at the same time" versions of DA_p and $\mathrm{MAP}_p^{\rhd}$. As is well known, the DA algorithm is invariant to specifications in which men propose individually (e.g. the lowest numbered unmatched man proposing in every round) instead of simultaneously. Similarly it is easily shown that $\mathrm{MAP}_p^{\rhd}$ is also invariant to such specifications. In some of our proofs we consider such a "one man propose at a time" version of $\mathrm{MAP}_p^{\rhd}$, assuming the reader's understanding that this does not change the outcome of the algorithm.

2.4. Independent Preferences

We primarily consider the case in which each man's value $u_m(w)$ for some $w \in \mathcal{W}$ is an i.i.d. draw from F_M, and each woman's value $u_w(m)$ for some $m \in \mathcal{M}$ is an i.i.d. draw from F_W. Throughout this section, a **random economy** refers to such independently drawn payoffs (for fixed M, W).

Our first result establishes a relationship between DA and a randomized version of MAP. Running DA_p on a random economy generates the same expected volume of marriages as would be obtained from running $\mathrm{MAP}_p^{\rhd}$, using a *uniformly randomly chosen* profile $\rhd$ of meeting orders. More generally, both methods lead to the same distribution over the number of marriages.[16]

Theorem 2.1 (random-order MAP gives the distribution of p-stable marriages)**.** *Fix sets M, W and prices $p = (p_m, p_w)$. Let K^{DA} be a random variable representing the number of p-stable marriages in a random economy. Let $K^{r\mathrm{MAP}}$ be a random variable representing the number of marriages created under $\mathrm{MAP}_p^{\rhd}$ for a random economy when each meeting order $\rhd_m$ is independently drawn from a uniform distribution over all orders. Then K^{DA} and $K^{r\mathrm{MAP}}$ have the same probability distribution.*

The distribution of K^{DA} obviously depends on prices p_M, p_W. However prices turn out to affect this distribution only to the extent that they affect the probability that any *one* man-woman pair are p-compatible. Recall that man m and woman w are p-compatible when $u_m(w) \geq p_m$ and $u_w(m) \geq p_w$. They are thus *incompatible* with probability

$$q(p_m, p_w) \equiv F_M(p_m) + F_W(p_w) - F_M(p_m)F_W(p_w) \tag{2.1}$$

We call $q(p)$ the **incompatibility parameter** at prices p.

[16] Any omitted proofs appear in Appendix B.

The next lemma states that, for fixed M, W, and meeting orders $(\rhd_m)_M$, the *expected* number of marriages created by $\mathrm{MAP}_p^{\rhd}$ at prices p is a function only of $q(p)$. In other words, if $q(p) = q(p')$ then $\mathrm{MAP}_p^{\rhd}$ and $\mathrm{MAP}_{p'}^{\rhd}$ induce the same expected number of marriages.

Lemma 2.1 (Expected MAP marriages is a polynomial function of q)**.** *Fix M, W and meeting orders $(\rhd_m)_M$. For any prices p_m, p_w, let $K_{p_m,p_w}^{\rhd}$ be a random variable representing the number of marriages created under* $\mathrm{MAP}(\rhd, p_m, p_w)$ *for a random economy. There exists a function $\bar{K}_{\rhd} : \mathbb{R} \to \mathbb{R}$ such that, for all p_m, p_w, $E[K_{p_m,p_w}^{\rhd}] = \bar{K}_{\rhd}(q(p_m, p_w))$, where $q(p_m, p_w) \equiv p_m + p_w - p_m p_w$ is the probability of incompatibility. Furthermore $\bar{K}_{\rhd}()$ is polynomial in its argument.*

This result extends to the expected number of p-stable marriages with the help of Theorem 2.1.

Theorem 2.2 (Expected p-stable marriages is a polynomial function of q)**.** *For any prices (p_M, p_W), let $K_{p_m,p_w}^{\mathrm{DA}}$ be a random variable representing the number of (p_m, p_w)-stable marriages for a random economy. Then $E[K_{p_m,p_w}^{\mathrm{DA}}]$ is a polynomial function of $q(p_m, p_w)$.*

Proof: Letting $\mathcal{O}$ denote the set of meeting order profiles, Theorem 2.1 implies

$$E[K_{p_m,p_w}^{\mathrm{DA}}] = E(K_{p_m,p_w}^{r\mathrm{MAP}}) = \frac{\sum_{\rhd \in \mathcal{O}} E[K_{p_m,p_w}^{\mathrm{MAP}(\rhd)}]}{|\mathcal{O}|}$$

where the numerator is a sum of polynomials in $q(p_M, p_W)$ by Lemma 2.1. $\square$

We now turn to our first main conclusion regarding the platform's revenue under p-stability. Namely, when the two sides' values are draw from the same distribution

$F = F_M = F_W$, the platform's expected revenue is symmetric in the price vector (p_M, p_W). We interpret this as a "non-price-discrimination" result: despite the fact that matched agents on the short side of the market extract greater value than those on the long side [e.g. Ashlagi et al. (2017)], the platform does not benefit from, for instance, specifically charging the short side a higher price.

Corollary 2.1 (Revenue symmetric in prices)**.** *Fix M and W, and suppose $F_M = F_W$. Let $R_{(p_M,p_W)}$ denote the platform's expected revenue from a random economy under a p-stable mechanism with prices $p = (p_M, p_W)$. Then we have $R_{(p_M,p_W)} = R_{(p_W,p_M)}$.*

The analogous result holds if the platform uses a $\mathrm{MAP}_p^{\triangleright}$ mechanism with arbitrary, fixed meeting orders $\triangleright$.

Proof: The planner earns (p_M+p_W) from each marriage, so under p-stability, expected revenue has the form $(p_M + p_W)\bar{K}(q(p_m, p_w))$ where $\bar{K}$ is some polynomial function of q from Theorem 2.2. Since $q()$ is symmetric in $F_M(p_M)$ and $F_W(p_W)$ the result follows. For $\mathrm{MAP}_p^{\triangleright}$, Lemma 2.1 leads to the analogous conclusion. □

This result demonstrates that a profit-maximizing platform does not price discriminate amongst the two sides of the market based on their relative sizes. When both sides' values are draw from the same distribution, if (p_M^*, p_W^*) is a revenue-maximizing price vector then so is (p_W^*, p_M^*).

In fact we now show that, under a standard assumption on the common distribution $F_M = F_W$, revenue maximizing prices are equal across the two sides, regardless of how unbalanced their sizes are. Of course that common price might change with respect to absolute market size or with the degree of market imbalance. Our point is that changes in market size or imbalance are not *per se* a reason for the platform to practice price discrimination.

The platform earns $p_m + p_w$ for each marriage. Fixing a DA (or MAP) algorithm, the price-optimization problem is thus

$$\max_{p_m,p_w}(p_m + p_w)E[K(p_m, p_w)] \tag{2.2}$$

$$= \max_{p_m,p_w}(p_m + p_w)\bar{K}(q(p_m, p_w)) \tag{2.3}$$

where $K(p_m, p_w)$ denotes the expected number of marriages created under the fixed algorithm for a random economy, which can be written in the form of $\bar{K}(q)$ by our two earlier results. Unsurprisingly, $\bar{K}(q)$ is decreasing in the incompatibility parameter q. For the case of DA, this statement is essentially identical to a Theorem of Gale and Sotomayor (1985b) stating that, under DA with *fixed* preferences, a preference truncation on one side of the market makes the other side worse off, weakly decreasing the number of marriages. Our proof (see Appendix B) also handles the case of MAP.

Lemma 2.2 (Monotonicity in q)**.** *For any* $\mathrm{MAP}_p^{\triangleright}$ *algorithm or any p-stable algorithm, the expected number of marriages $\bar{K}(q)$ as a function of incompatibility q is decreasing in q.*

Now if we commit the platform to any fixed incompatibility rate $q = q(p_m, p_w)$, the platform would prefer to maximize $p_m + p_w$ subject to that constraint. Under a standard hazard rate assumption, this will imply $p_m = p_w$. When preferences for both sides are drawn from the same distribution, the platform charges both sides the same price regardless of their relative sizes.[17]

Theorem 2.3 (Monotone hazard rate implies equal pricing)**.** *Fix M and W, and suppose $F = F_M = F_W$, where F has a weakly increasing hazard rate $\frac{f(u)}{1-F(u)}$. Let $R_{(p_M, p_W)}$ denote the platform's expected revenue from a random economy under a p-stable mechanism with prices $p = (p_M, p_W)$. If (p_M^*, p_W^*) maximizes $R_{(p_M, p_W)}$ then $p_M^* = p_W^*$.*

The analogous result holds if the platform uses a $\mathrm{MAP}_p^{\triangleright}$ mechanism with arbitrary, fixed meeting orders $\triangleright$.

Without the monotone hazard rate condition, equal pricing may not be optimal. To illustrate this in as simple a way as possible, consider the following discrete example.[18]

Example 2.1 (Optimal, unequal prices)**.** *Consider one man and one woman. Each (independently) has a value of being matched to the other that is either 0.1 (with probability π) or 0.9 (probability $1-\pi$). We can restriction attention to prices $(p_m, p_w) \in \{0.1, 0.9\}^2$. Depending on π, one can easily check that the following price combinations maximize expected revenue.*

$$
\begin{aligned}
(p_m^*, p_w^*) &= (0.9, 0.9) && \textit{when } \pi \leq 4/9, \\
(p_m^*, p_w^*) &\in \{(0.1, 0.9), (0.9, 0.1)\} && \textit{when } 4/9 \leq \pi \leq 4/5, \\
(p_m^*, p_w^*) &= (0.1, 0.1) && \textit{when } 4/5 \leq \pi.
\end{aligned}
$$

---

[17] Obviously the result assumes that both sides' values are drawn from the same distribution; relaxing such an assumption clearly gives the platform incentive to charge higher prices to the side of the agent which tends to captures higher values from being matched.

[18] Though the discrete example does not fit in our domain of analysis, one can (tediously) perturb it into a continuous version that yields the same conclusion.

The case $4/9 \leq \pi \leq 4/5$ *is the relevant one, demonstrating unequal optimal prices. However the* set *of optimal price lists is symmetric, in accordance with Corollary 2.1.*

2.4.1. Constrained Serial Dictatorship: MAP with Equal Meeting Orders

Our next objective is to evaluate the platform's expected revenue from any pair of prices $p = (p_m, p_w)$ and value distributions F_M, F_W. When the platform uses the DA_p mechanism, a closed-form expression for revenue remains elusive. On the other hand, a platform's expected revenue from DA_p turns out to be remarkably close to that from using $\mathrm{MAP}_p^{\rhd}$ (with the same prices p), when *all men's meeting orders are identical* under $\rhd$.

When $\rhd$ is such that all men have the same meeting order, $\mathrm{MAP}^{\rhd}$ can be viewed as a *constrained* serial dictatorship mechanism: in turn, each woman takes her turn choosing her favorite man among the remaining men *who are compatible with her* (if any).[19] For this subclass of MAP algorithms, the probability distribution of the number of marriages can be described explicitly. To do this, we introduce the concept of *q-analogs* (parameterized generalizations) of integers, factorials, and binomial coefficients.

Definition 2.6. *For any real number* $q \in [0, 1]$*, the* ***q-analog of integer j*** $\in \mathbb{Z}$ *is*

$$[j]_q \equiv 1 + q + \cdots + q^{j-1} = \frac{1 - q^j}{1 - q}$$

[19] Though it is not our objective, one can imagine this mechanism describing a platform where agents on one side (women) arrive randomly, and get to choose a among all remaining (but compatible) agents on the other side.

*The **q-factorial of j** is*

$$[j]_q! \equiv [j]_q[j-1]_q \cdots [1]_q$$

*The q-**binomial coefficient** for integers $k, n \in \mathbb{Z}_+$, with $k \le n$, is*

$$\begin{bmatrix} n \\ k \end{bmatrix}_q \equiv \frac{[n]_q!}{[k]_q![n-k]_q!} = \frac{(1-q^n)(1-q^{n-1})\cdots(1-q^{n-(k-1)})}{(1-q^1)(1-q^2)\cdots(1-q^k)}$$

The following result establishes the probability distribution of the number of marriages under an equal-meeting-order MAP algorithm in terms of these q-analogs.[20]

Theorem 2.4 (Distribution of marriages for identical orders)**.** *Fix M, W, and a profile of meeting orders $\rhd$ in which all men have identical meeting orders: for all $m, m' \in M$, $\rhd_m = \rhd_{m'}$. Fix prices p_M, p_W and incompatibility parameter $q = q(p_M, p_W)$. Let $K^=$ be a random variable representing the number of marriages in a random economy created under $\mathrm{MAP}_p^{\rhd}$. The probability distribution of $K^=$ is given by*

$$P(k; M, W) = (1-q)^k q^{(M-k)(W-k)} \begin{bmatrix} M \\ k \end{bmatrix}_q \begin{bmatrix} W \\ k \end{bmatrix}_q [k]_q! \tag{2.4}$$

The first two terms in Equation 2.4 have a straightforward interpretation: $(1-q)^k$ is the probability of mutual compatibility among k men matched to k women, and $q^{(M-k)(W-k)}$ is the probability of mutual incompatibility among all possible pairs of the remaining

[20] The distribution of Equation 2.4, called the absorption distribution, was described by Blomqvist (1952). Kemp (1998) finds its moments and shows that it is log-concave. Ebrahimy and Shimer (2010) make use of this distribution.

agents. The remaining q-analog terms are a probabilistic analog to $\binom{M}{k}\binom{W}{k}k!$, the number of possible size-k matchings in a market (M, W).

Theorem 2.4 is itself a q-analog to the two trivial observations that (i) at $q = 1$ the probability of zero matches is one, and (ii) at $q = 0$ the probability of $k = \min\{M, W\}$ matches is one. Typically, of course, a revenue maximizer would indirectly set an intermediate value of q (through its choice of prices p_m, p_w), and might care only about the expectation of $K^=$. Kemp (1998) provides the following expression.

Theorem 2.5 (Kemp, 1998)**.** *The expected value of $K^=$ (defined in Theorem 2.4) is*

$$E(K^=) = \sum_{j=1}^{\min\{M,W\}} \frac{[(1-q^M)\cdots(1-q^{M-j+1})][(1-q^W)\cdots(1-q^{W-j+1})]}{1-q^j} \tag{2.5}$$

Figure 2.1 illustrates Equation 2.5 for a fixed number of men $M = 50$, varying both W and the incompatibility parameter q, where values are assumed to be uniform on $[0, 1]$. The figure highlights the obvious fact that, as W grows large, the expected number of marriages converges to 50: when the market is very imbalanced, the platform can charge a very high price and still create the close to the maximum number of possible marriages. Also obvious is that, when prices are very low, the platform again creates close to the maximum number of possible marriages, regardless of the market imbalance. On the other hand, when prices are high and the market is not very imbalanced, a non-trivial number of agents are unmatched. In these situations the platform faces a less trivial pricing decision.

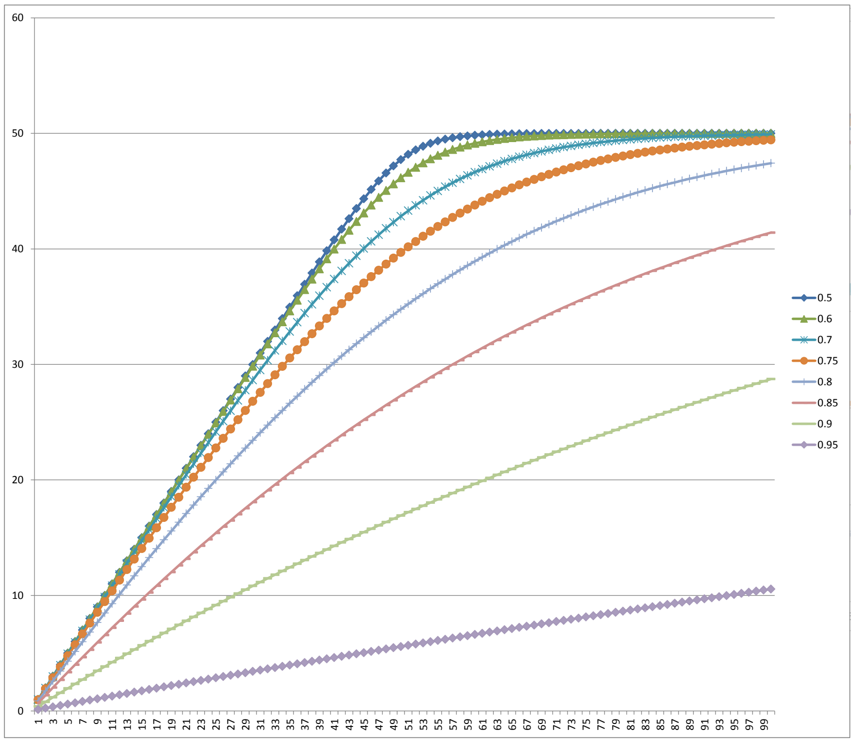

Figure 2.1. Expected marriages under equal-order MAP. Fixing $M = 50$ men and varying $1 \leq W \leq 100$ women, the figure shows the expected number of marriages for various values of q. For example, when $q = 0.9$, if $M = W = 50$, Equation 2.5 yields roughly 43.6 expected marriages. If values $u_m(\cdot), u_w(\cdot)$ were uniformly i.i.d. on $[0, 1]$, this value of q corresponds to prices of approximately $p_m = p_w = 0.684$.

2.4.2. Constrained Serial Dictatorship as a Bound

Returning to our main motivation of p-stability, we observe that for relatively smaller markets, the expected number of p-stable marriages is well-approximated by Equation 2.5. This would be less surprising in relatively unbalanced markets, where many mechanisms might yield an expected number of marriages close to $\min\{M, W\}$. Therefore we specifically consider balanced markets in Figure 2.2, where we graphically compare the

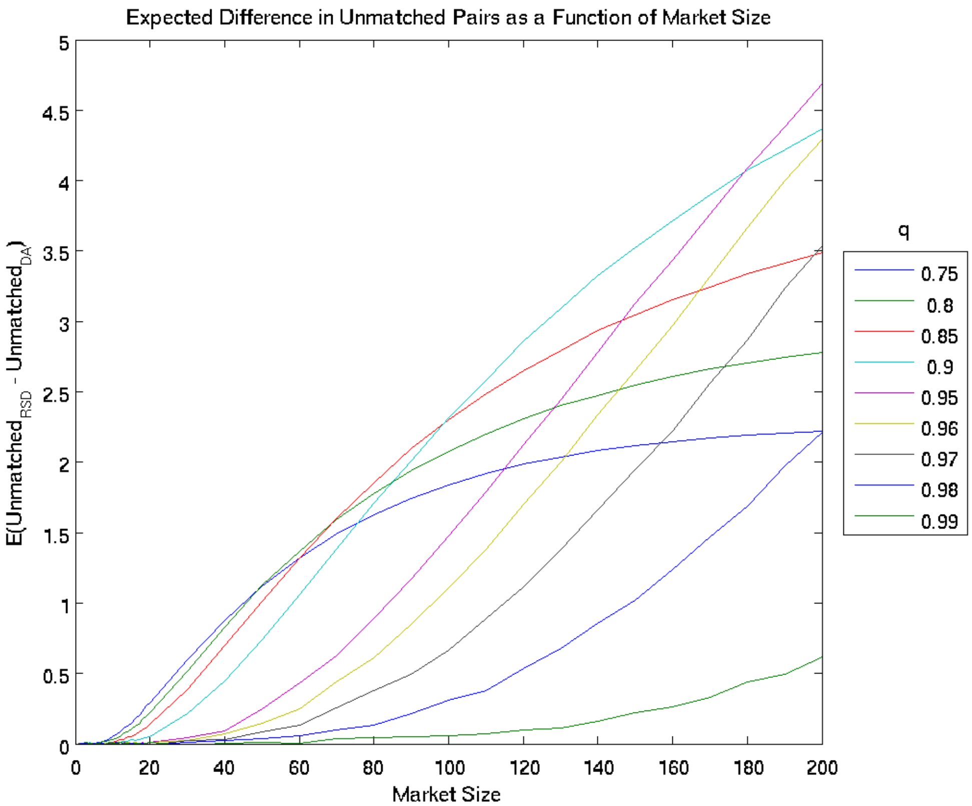

Figure 2.2. For balanced markets and various values of q, the graph shows the difference between Equation 2.5 and the expected number of p-stable marriages. Hence the former is a good approximation for the latter in small, balanced markets. The fact that the values are all positive is consistent with a related result of Arnosti (2016).

(simulated) expected number of p-stable marriages to the expected value under Equation 2.5. It is in balanced markets that one might expect the largest divergence between these two values. As one can see, Equation 2.5 provides a fairly good approximation of the expected number of p-stable marriages for the market sizes considered in the figure.

Figure 2.2 additionally suggests that the expected difference in the number of unmatched pairs between the constrained serial dictatorship and DA stabilizes as the market grows large. This means that the *fraction* matched under the constrained serial dictatorship converges to the fraction matched under DA. Thus the approximation becomes arbitrarily accurate for the pricing purposes of the platform in large markets.

Therefore we turn to the analysis of the expected number of matches in balanced large markets, i.e. where $M = W = n$ for some n. Fixing the incompatibility parameter q, Equation 2.4 gives us the following probability that MAP$^{\triangleright}$ yields a *perfect match* (n marriages) when all men have the same meeting order.

$$\begin{aligned} P(n;n,n) &= (1-q)^n q^{(n-n)(n-n)} \begin{bmatrix} n \\ n \end{bmatrix}_q \begin{bmatrix} n \\ n \end{bmatrix}_q [n]_q! \\ &= (1-q)^n [n]_q! \\ &= (1-q)^n \frac{(1-q^n)\cdots(1-q^1)}{(1-q)^n} \\ &= (1-q^n)\cdots(1-q^1) \end{aligned}$$

As $n \to \infty$, $P(n;n,n)$ converges to the following expression.[21]

$$\phi(q) \equiv \prod_{i=1}^{\infty}(1-q^i) \tag{2.6}$$

Since $\phi(q)$ is positive whenever $q < 1$, for any prices $p_m, p_w < 1$ (i.e. $F(p_m), F(p_w) < 1$), the probability that equal-order MAP yields a perfect match remains bounded away from zero as the market grows arbitrarily large.

[21] This expression has been referred to as *Euler's function*, though other functions also carry that name. It is a special case of the q-Pochhammer symbol. Banerjee and Wilkerson (2016) give a closed-form approximation for ϕ.

To generalize this observation, consider the asymptotic probability that any fixed number of men and women remain unmatched. If there are k marriages among M men and W women, then there must be $g = M - k$ single men and $h = W - k$ single women. Rewriting Equation 2.4 in terms of g and h we have

$$P(k; M, W) = P(k; k+g, k+h) = (1-q)^k q^{gh} \begin{bmatrix} k+g \\ k \end{bmatrix}_q \begin{bmatrix} k+h \\ k \end{bmatrix}_q [k]_q! \tag{2.7}$$

Now fix a constant number of single men and women, g and h respectively, and let the number of couples $k \to \infty$ (so implicitly, $M = k + g \to \infty$ and $W = k + h \to \infty$). Then Equation 2.7 converges to the following (see Appendix B for proof).

$$\lim_{k\to\infty} P(k; k+g, k+h) = \phi(q) \frac{q^{gh}}{(1-q)\cdots(1-q^g)\cdot(1-q)\cdots(1-q^h)} \tag{2.8}$$

For the case of balanced markets (where $M = W$, so $g = h$) this gives us the following.

Theorem 2.6 (Asymptotic distribution of unmatched agents in balanced markets)**.** *Fix prices p_m, p_w with incompatibility parameter $q \equiv q(p_m, p_w) < 1$. Consider running equal-order* MAP *(all men have identical meeting orders) on a random economy. For any positive integer $g \in \mathbb{Z}_{++}$,*

$$\begin{aligned} &\lim_{M=W\to\infty} P(\textit{exactly } g \textit{ men remain single}) \\ &= G(g; q) \equiv \phi(q) \frac{q^{g^2}}{((1-q)\cdots(1-q^g))^2} \end{aligned} \tag{2.9}$$

For the case $g = 0$, the limit is $G(0; q) = \phi(q)$.

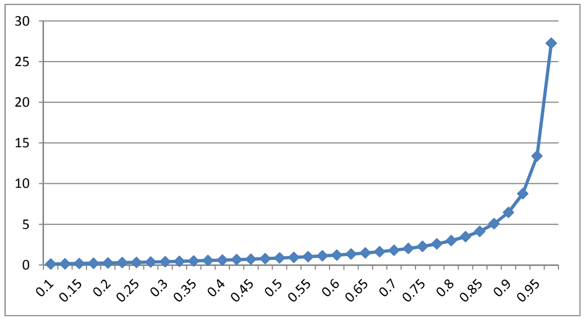

Figure 2.3. The expected number of unmatched agents under equal-order $\text{MAP}_p^{\triangleright}$ in large, balanced markets, for incompatibility parameter $0 \leq q \leq 0.975$.

Equation 2.9 can be used to approximate the *expected* number of unmatched agents in large (balanced) markets under equal-order MAP.[22] Figure 2.3 shows the expected number of single men (and single women) in balanced markets as $n = M = W$ grows arbitrarily large. For example, even when any pair is incompatible with probability $q = 0.95$, the expected number of single men in a large balanced market is less than 13.4. For $q = 0.80$ the value drops below 3.

Since single agents are a vanishingly small fraction of the total market for any q, a platform using equal-order MAP could charge close to maximal prices to both sides of the market and still create close to the maximal feasible number of matches; i.e. revenue-maximizing prices converge towards $p_m = p_w = 1$ as the market grows large. For example, in a balanced market with i.i.d. preferences from $U[0,1]$ and 100 agents on each side, the platform optimally charges a symmetric price of 0.77; with the same preferences and 500 agents on each side, the platform optimally charges a symmetric price of 0.86.

[22] The probabilities $G(g;q)$ quickly converge to zero as g increases, so we approximate the asymptotic expectation $\sum_{g=0}^{\infty} gG(g;q)$ by considering only the first (sufficiently many) terms.

The more surprising fact is the strength of the approximation for moderately-sized markets.[23] In particular, Figures 2.2 and 2.3 jointly suggest that the difference in the expected number of matches in any size of market markets is bounded above by the asymptotic number in large markets *for any* incompatibility parameter q. Furthermore, Figure 2.3 shows that this asymptotic number is small for "most" values of q.

The simulations suggest two further results for p-stable matchings that we have been unable to prove analytically. First, Equation 2.5 appears to be a lower bound on the expected number of p-stable marriages for all market sizes. This is strongly suggested by the non-negativity in Figure 2.2 across a range of market sizes and probabilities of incompatibility. Additionally this is consistent with a result by Arnosti (2016) in a similar setting. Second, the expected number of single agents under p-stability appears to converges to zero. This would imply that the "price of stability" is zero in the limit.[24]

2.5. Correlated Preferences

Correlation among user preferences in the one-to-one matching case introduces incentives for the platform to price-discriminate based on relative market size. This contrasts with the literature on two-sided markets, where correlation does not affect the prices of a risk-neutral platform.[25] In particular, in the absence of capacity constraints, the many-to-many platform optimally sets prices *as if* it could optimally price each transaction.

[23] Such markets may also be of most practical relevance to the platform. Consider for example Airbnb or Uber. The gross number of participating agents is large, but most agents are matching in much smaller localized markets, both geographically and temporally. It is plausible that many such localized markets would fall into the intermediate range.

[24] From Erdos and Renyi (1964), we already know that the "price of IR" is zero in our model of i.i.d. preferences: as the market size grows, it is with probability approaching 1 that there exists a perfect matching satisfying individual rationality. The additional observation here is that the pairwise-blocking constraint is also asymptotically costless.

[25] E.g. Rochet and Tirole (2003), Armstrong (2006), Rochet and Tirole (2006).

This obviates the need for the platform to consider correlation. In contrast, one-to-one matching introduces competition *within* the platform. Agents can only be matched once. Therefore each match has an opportunity cost, and this opportunity cost depends on the correlation structure.

Two-sided matching naturally suggests two dimensions of correlation. Consider the preferences of women over men. First, women may generally agree in their rankings over men. In this case we say preferences are correlated *across* the women. Second, each woman's valuations over the men may be correlated. This occurs naturally if the women have heterogeneous outside options, for example. In this case we say preferences are correlated *within* the women.

The type of correlation matters for pricing in the one-to-one matching case. Assume a continuum of agents to abstract from matching frictions. Since we are interested in how relative market size affects pricing, consider two sides of the market: a long side L with mass 1, and a short side S with mass $\lambda < 1$. For simplicity of exposition we assume that individual match values are identically distributed on the two sides: $F_M = F_W = F$. Finally, we examine relative prices in the corner cases where preferences are either perfectly correlated across or within.

In particular, *relative* pricing implications across the two sides of the market are opposite for the two types of correlation. When preferences are correlated across, the long side is charged a lower price because the marginal agent on the long side receives a higher-quality match. When preferences are correlated within, the long side is charged a higher price because the value, or outside option, of the marginal agent on the long side is higher.

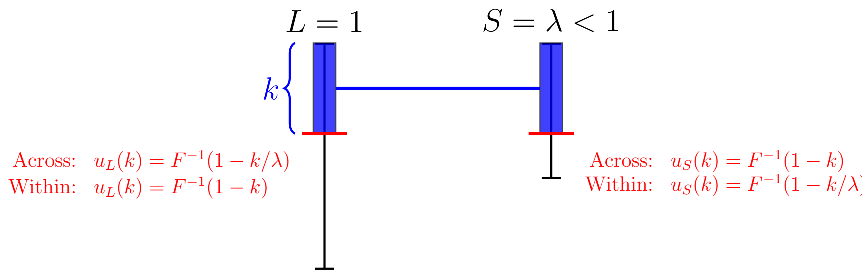

Figure 2.4. When preferences are correlated across, the marginal valuation is higher on the short side, implying $p_L < p_S$. When preferences are correlated within, the marginal valuation is higher on the long side, implying $p_L > p_S$.

Graphical intuition for relative pricing in the two cases of correlation is provided in Figure 2.4. First, note that in either case the platform maximizes revenue at some pair of prices where the masses of agents willing to pay are equal on both sides.[26] When preferences are perfectly correlated across, the match is assortative. If a mass k of agents is matched, the value to the marginal agent on the long side is the match value of the k-th quantile agent on the short side, given by $F^{-1}(1 - k/\lambda)$. When preferences are perfectly correlated within, the agents are indifferent over each agent on the other side of the market. Therefore the value to the marginal agent on the long side is the outside option, given by $F^{-1}(1 - k)$. Arguments are symmetric for the marginal valuations on the short side of the market. Note that in the across case, the value of the marginal agent is given by the marginal agent on the *other* side; in the within case, the value is given by the marginal agent on *same* side.

The dependence of relative pricing on correlation in the one-to-one matching case suggests the pricing matrix presented in Figure 2.5. The pricing matrix considers relative pricing implications across two dimensions: preference correlation and *intra*-platform

[26] If, for example, there are more acceptable men than women, the platform can raise the women's price without reducing the number of resulting marriages.

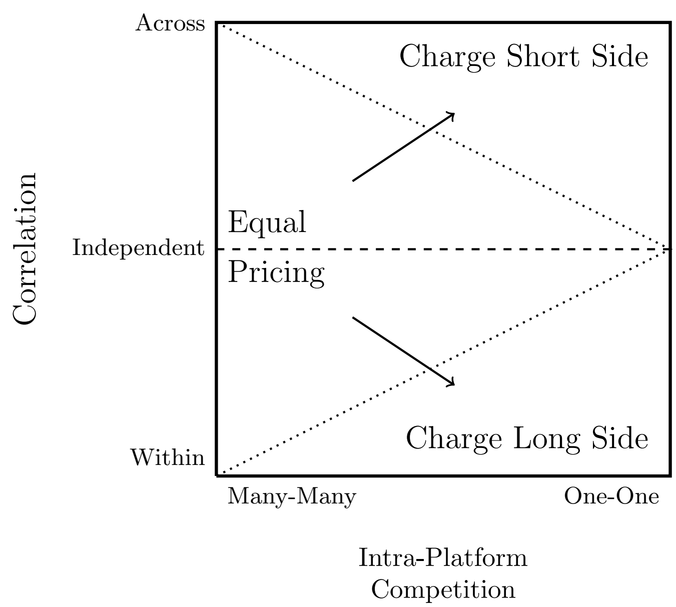

Figure 2.5. This figure captures when and how correlation moves the platform away from equal pricing. In the absence of intra-platform competition (many-to-many case), correlation does not affect the equal pricing recommendation. With extreme intra-platform competition (one-to-one case), any correlation affects the equal pricing recommendation. Independent preferences are a knife-edge case.

competition. On the vertical axis, the correlation of agent preferences may vary.[27] On the horizontal axis, the platform may face different levels of intra-platform competition as a result of matching capacity constraints. The many-to-many case presents one extreme with no capacity constraints, and hence no intra-platform competition. The one-to-one case presents the other extreme, where each agent faces a capacity constraint of one.

The pricing matrix summarizes our key insights on the relationship between relative pricing, preference correlation, and intra-platform competition (capacity constraints). First, when preferences are independent both across and within agents, a monopolist

[27] Note that both types of correlation will generally exist in agent preferences, in which case we hypothesize that each type of correlation will exert the pressure described, with a magnitude commensurate to the extent of correlation.

platform optimally charges equal prices for equal match-value distributions, regardless of capacity constraints. Second, in the absence of capacity constraints, the platform charges equal prices regardless of correlation structure. Third, equal pricing becomes a knife-edge case with one-to-one capacity constraints: relative pricing generally depends on market size and depends crucially on the type of correlation. We hypothesize that these insights vary “continuously” with the extent of capacity constraints and the correlation structures present in the market. However, we leave any formalization of this argument to future work.

A final comparative static of interest is the effect of changes in the relative market size, captured by the parameter λ. This occurs with ride-sharing platforms, for example, when the pool of potential passengers suddenly experiences a “surge” due to a special event or other demand shock. In particular, preferences on a ride-sharing platform exhibit significant within correlation. Passengers have different values of hailing a cab but cannot choose the driver. Similarly, drivers care about picking up a fare but are largely indifferent about the passenger.[28] In the case of within correlation, it is possible to show that an increase in λ leads to higher prices for the “surging” side. This is consistent with the notion of “surge-pricing” practiced by ride-sharing services such as Uber. Interestingly, however, in situations where preferences are primarily correlated across, the model suggests the opposite pricing implication.

[28] While passenger and driver preferences also exhibit correlation across, we believe this is a second-order consideration for the pricing decisions of a monopolist platform.

2.6. Conclusion

The proliferation of online platforms and their propensity toward increased market concentration have intensified interest in the study of platform pricing. Two features – exclusivity and heterogeneity – jointly characterize a number of important platforms, yet this combination of features remains largely unaddressed by the literature on two-sided markets. Additionally, interaction between the literatures in two-sided markets and stable matching remains limited, despite intriguing similarities. We address both gaps by studying the revenue maximization problem of a monopolistic matching platform that sets prices to two sides of a "marriage market."

Our first pricing result is that a platform does *not* want to price discriminate across the two sides of the market based on their relative sizes when agent preferences are independently drawn. This result appears counterintuitive in light of a result by Ashlagi et al. (2017), who show that small imbalances in size create large imbalances in surplus across the two market sides. Our proof technique introduces a class of "meet-and-propose" (MAP) matching algorithms that generalize DA by relaxing the equivalence between preferences and meeting orders. Even though MAP algorithms do not guarantee stable outcomes, the expected number of matches under a randomized MAP algorithm coincides with the expected number of stable matches.

Next we address the problem of which prices a stability-constrained platform should optimally charge. Our recursive formulation for the expected number of stable matches quickly becomes intractable with market size. In the absence of a closed-form solution for revenue under stability, we therefore show via simulations that the revenue under stability is well-approximated using an *exact* expression for revenue under a constrained serial dictatorship mechanism. While the approximation is good for large markets, this is driven largely by the vanishing cost of stability. In this case the platform can extract almost all surplus under either mechanism. We therefore place particular emphasis on the fact that the the approximation remains good for moderately sized markets. We suggest that this intermediate category may also be the most relevant in practice, given that many large markets are locally small.

Our findings depart significantly from the literature on two-sided markets when we consider correlated preferences. In particular, correlation among user preferences in the one-to-one matching case introduces incentives for the platform to price-discriminate based on relative market size. In contrast, correlation does not affect optimal pricing in the many-to-many setting of canonical two-sided markets. Furthermore, the type of correlation matters for one-to-one matching. Preferences that are correlated across a group of agents have the opposite pricing implication as preferences that are correlated within agents. Our analysis of pricing under perfect correlation within and across agents suggests that relative market size matters in the presence of capacity constraints, which introduce intraplatform competition. We leave this topic as an interesting avenue for future research.

CHAPTER 3

A Dynamic Theory of Political Slant in News Media (joint with Tomer Hamami)

3.1. Introduction

The literature on political slant in news media is relatively young. Models in this subject area are overwhelmingly static in nature. The biased beliefs of consumers are assumed to be exogenous, and news firms simply slant their news in order to meet demand. Such models overlook the critical role news firms play in the formation of consumer beliefs in the first place. The purpose of this paper is to analyze this role in a dynamic setting where news slant affects consumer beliefs over time.

We extend the seminal model of Mullainathan and Shleifer (2005; hereafter MS (2005)) to a dynamic setting. In the original model, consumers have (exogenous) biased beliefs about the state of the world and a taste for news slanted towards their beliefs (to provide tension in the model, consumers also dislike slanted news in general). MS (2005) show that a profit-maximizing monopolist slants towards the consumers' bias, but profit-maximizing duopolists actually slant toward a point even more extreme than the most biased consumer's belief. This result follows from standard Hotelling reasoning: duopolists position themselves farther apart than otherwise necessary to mitigate the effects of competition on price.

In that paper, a monopolist's profits are higher when consumers less biased, but duopolist profits are higher when consumers are more biased (to a certain point). It follows that, if news firms can influence consumer beliefs, a monopolist would want to manipulate consumers towards a lack of bias, while a duopolist would profit from some degree of polarization. This idea forms the crux of this essay – while other papers frame bias in news media as firms simply responding to consumer demand, the implications here are more insidious: the news media has incentive to create biased consumers on its own, even if consumers have initially unbiased beliefs.

We construct a dynamic model where the MS (2005) game is the stage game. Additionally, news firms can choose not to report news in a period after observing the data available to report. Consumers are naive about nondisclosure; they do not update beliefs after observing no news even though the absence of reporting is informative about the state of the world. Firms can use this disclosure/nondisclosure decision to influence consumer beliefs. Although a news firm forgoes current period revenue by not reporting, the next period state may be more profitable following nondisclosure.

We first solve the problem for the monopolist news firm, which effectively reduces to a dynamic programming problem. We prove that the optimal disclosure set for the monopolist results in consumer beliefs converging to unbiasedness. Of course, this finding depends on the monopolist being motivated solely by profit; the journalism industry is competitive in part because this assumption about a monopolist news firm may not be true. Simulations and computational analysis corroborate this finding.

We next analyze the far more complicated duopolist problem. We show the incentive to push consumers towards unbiasedness is tempered by the incentive to maintain consumer differentiation in order to suppress price competition. Moreover, we show how news firms have incentive to actually create political division even in the case where consumers were originally homogeneous and unbiased. Simulations and computational analysis demonstrate scenarios in which this occurs.

The remainder of this paper is organized as follows: In the rest of this section, we discuss what is meant by "political slant" and review the literature on bias in news media. In Section 3.2, we set up the theoretical framework of MS (2005) and show how we extend it to a dynamic setting. In Section 3.3, we analyze the monopolist's problem. In Section 3.4, we analyze the duopolist's problem. We conclude in Section 3.5. All proofs and figures are relegated to the appendix.

What is Political Slant? Our usage of the term political slant is consistent with the meaning common in the literature. In particular, slant does not mean the transmission of factually incorrect information (i.e. "fake news"). Rather, slant involves a combination of framing, phrasing, and filtering to present information in a way that makes it seem as if its implications are consistent with a certain set of preexisting beliefs. To demonstrate, we gently borrow the below example from Gentzkow and Shapiro (2006):

> "On December 2, 2003, American troops fought a battle in the Iraqi city of Samarra. Fox News began its story on the event with the following paragraph:

> In one of the deadliest reported firefights in Iraq since the fall of Saddam Hussein's regime, U.S. forces killed at least 54 Iraqis and captured eight others while fending off simultaneous convoy ambushes Sunday in the northern city of Samarra.

The *New York Times* article on the same event began:

> American commanders vowed Monday that the killing of as many as 54 insurgents in this central Iraqi town would serve as a lesson to those fighting the United States, but Iraqis disputed the death toll and said anger against America would only rise.

And the English-language website of the satellite network Al Jazeera (Al-Jazeera.net) began:

> The US military has vowed to continue aggressive tactics after saying it killed 54 Iraqis following an ambush, but commanders admitted they had no proof to back up their claims. The only corpses at Samarra's hospital were those of civilians, including two elderly Iranian visitors and a child."

These three news stories each reported on the same event, and none said anything untrue. Nevertheless, it is clear that each paints a very different picture about what occurred and the implications of the event. This is what is meant by political slant.

Related Literature. The literature on political slant in news media is largely a post-internet phenomenon and thus still a relatively young body of scholarship. Gentzkow et al. (2015) do an excellent job of breaking down existing theoretical work into demand-side and supply-side models of bias. A demand-side model of media bias is one where consumers prefer to consume biased news, whether the source of that preference is explicit bias [Mullainathan and Shleifer (2005), Gentzkow and Shapiro (2006), Burke (2008), Stone (2011)] or simply differentiation according to partisanship [Bernhardt et al. (2008), Chan and Suen (2008)]. Whichever assumption drives consumer preferences for slanted news, a key feature of all of these papers is that it is exogenous, and news firms are merely responding to a force that already exists. As a direct extension of MS (2005), the current paper falls into this category, and its key contribution is to examine the mechanism by which consumer preference for bias is endogenously affected by the actions of news firms themselves. Supply-side models of bias, on the other hand, are ones where slanted news arises not because of consumer preference, but because either the news firm (or its agents) wants to affect policy [Baron (2006), Chan and Suen (2009), Duggan and Martinelli (2011), Anderson and McLaren (2012)], or because political agents themselves are manipulating voting behavior using the media as a catalyst [Strömberg (2004)].

There is also a body of work that uncovers empirical evidence that political slant actually exists and varies across news outlets [(Groseclose and Milyo (2005), Gentzkow and Shapiro (2010)]. Gentzkow and Shapiro (2010) is especially relevant in that they find evidence consistent with the theory that news slant is primarily driven by demand-side forces. Finally, there are a few papers that examine the extent to which slanted news media can impact beliefs and/or political behavior [Druckman and Parkin (2005), Dellavigna and Kaplan (2007), Gerber et al. (2009), Hmielowski et al. (2015), Martin and Yurukoglu (2017)]. Each finds evidence that such influence exists, with a particular emphasis on existence and availability of the Fox News network.

3.2. Framework

There is some state of the world ω that is distributed normally with mean 0 and variance v_ω. Each period, the firm observes a data point $d_t = \omega + \epsilon$ where ϵ is normally distributed with mean 0 and variance v_ϵ. A news firm can observe data point d_t and then choose to report news $n \in R$. In this context, we define news slant as $s \equiv n - d$. A news firm chooses a one-to-one slanting strategy $s : R \to R$ that maps each possible data point to a unique news report. This restriction on the firm's strategy space is intended to represent some required standard of editorial conduct; the firm can slant news in a way that is pleasing to its readers, but it cannot explicitly obfuscate a news event. A firm's disclosure set $D \subseteq R$ is the set of data points for which a firm decides to report news.

A unit mass of consumers have potentially biased beliefs prior to observing any news. Specifically, they initially believe ω is normally distributed with mean b_1 and variance v_ω. Consumers have the following preferences regarding news:

- Consumers are inherently interested in learning about ω and thus value learning the realization of d each period.
- Consumers dislike slanted news in general. This could be true because of the psychological burden associated with unraveling the true value of d from some $n \neq d$ or because slanted news is in general of lower journalistic quality.

- Consumers prefer news to be slanted in accordance with their existing biases, represented in period t by b_t.

Thus, if a consumer with bias b consumes news n with slant s, his payoff is characterized as follows:

$$U_b = \bar{u} - \chi s^2 - \phi(n - b)^2 - P \tag{3.1}$$

Where $\bar{u}$, χ, and ϕ are positive parameters. In this expression, $\bar{u}$ represents the inherent value of information, χ represents the consumer's general distaste for slant, ϕ represents the consumer's distaste for news that fails to conform to his preexisting biases, and P is the price the consumer pays for the news. Each consumer purchases news from the news firm for which its expected payoff is highest and also has an outside option with value equal to zero. The news firm's payoff is simply equal to its revenue, which is equal to the price it charges (P) multiplied by the share of the market captured, and discounts future payoffs exponentially with discount factor δ.

The key distinction between this model and that of MS (2005) is that this game is played repeatedly, with a new data draw each period, and consumer beliefs b_t evolve over time. A consumer that consumes news in period t effectively observes data point d_t and thus updates his beliefs as follows:

$$b_{t+1} = \frac{p_t b_t + p_\epsilon d_t}{p_t + p_\epsilon} \tag{3.2}$$

$$p_{t+1} = p_t + p_\epsilon \tag{3.3}$$

Where $p_t \equiv \frac{1}{v_t}$ and $p_\epsilon \equiv \frac{1}{v_\epsilon}$ are the corresponding precisions. However, consumers are boundedly rational in the following sense:

Assumption 3.1. *Consumers behave as though the firm never withholds news. That is, if the firm reports no news, then consumers do not update their beliefs even though the absence of disclosure may imply something about d_t. Moreover, when a consumer considers his expected payoff when determining his willingness to pay, he takes expectations over the full distribution of d_t even if the firm might not report news for all d_t.*

There is a growing body of empirical and experimental literature that suggests consumers failing to update in the absence of information may actually be a more accurate representation of human behavior than the fully rational consumer. For example, Jin et al. (2015) show in an experimental setting that information receivers are insufficiently skeptical of nondisclosure in a verifiable disclosure game. Similarly, Brown et al. (2012) show that consumers fail to infer low quality about movies whose studios choose not to make them available to critics prior to release.

This assumption allows us to exploit the conjugate prior for the normal distribution and characterize consumer beliefs in any period t. Let 1_{D_t} be an indicator function that is equal to 1 if $d_t \in D_t$ and 0 otherwise. Then, for any period $t \geq 1$, consumer beliefs are distributed normally with mean b_t and variance v_t as follows:

$$b_t = \frac{p_\omega b_1 + p_\epsilon \sum_{\tau=1}^{t-1} d_\tau 1_{D_\tau}}{p_\omega + p_\epsilon \sum_{\tau=1}^{t-1} 1_{D_\tau}} \tag{3.4}$$

$$v_t = \frac{1}{p_t} = \frac{1}{p_\omega + p_\epsilon \sum_{\tau=1}^{t-1} 1_{D_\tau}} \tag{3.5}$$

Formally, the timing of the game is as follows:

(1) The firm chooses slanting policy $s_t(d_t)$. This policy is common knowledge. If there are multiple firms, they choose their policies simultaneously.

(2) The firm chooses a price P_t to charge if it reports news. This price is common knowledge. If there are multiple firms, they choose their prices simultaneously.

(3) Consumers decide whether or not to purchase from the firm (conditional on news being reported). If there are two firms, each consumer purchases from at most one firm.

(4) The firm privately observes d_t. If there are two firms, each observes the same d_t.

(5) The firm decides whether or not to report news. If a firm reports, then $n_t = d_t + s_t(d_t)$. Only consumers who previously decided to purchase news from this firm know whether or not news was reported.

(6) Stage game payoffs are realized. Any consumer that observes news updates his beliefs accordingly, and the game returns to Step 1 for period $t + 1$.

Before we proceed to analyze the problems of monopolist and duopolist news firms, it will be useful establish some shorthand notation:

$$\gamma \equiv \frac{\phi}{\chi + \phi} \tag{3.6}$$

$$s^z(d) \equiv \gamma(z - d) \tag{3.7}$$

The parameter γ will frequently recur as a coefficient. It measures, on a scale from 0 to 1, the extent to which consumers are willing tolerate slanted news in order to consume news that conforms to their current beliefs. MS (2005) show that $s^z(d)$ is the slanting policy that maximizes the expected payoff of a consumer with belief $b = z$. In our solutions, firms will be exclusively using slanting policies with this structure, so the shorthand is useful. We can also say that a news firm that chooses slanting policy $s^z(d)$ is positioning itself at location z in a spatial modeling sense. Lastly, it will be useful to note that the expected payoff for a consumer that consumes news positioned at z and pays price P is:

$$E_d[U_b] = \bar{u} - \chi\gamma(b^2 + v_d) - \phi\gamma(z - b)^2 - P \tag{3.8}$$

For the derivation of this expression, see Lemma A1 from MS (2005).

3.3. Monopolist's Problem

3.3.1. Monopoly: Theory

In this section, we will assume that consumers are homogeneous and have beliefs (b_t, p_t) in period t. [1] The strategy for the monopolist in period t consists of a slanting policy $s_t(d_t)$, a price P_t, and a disclosure set D_t. The strategy for a consumer in this case is simply whether or not to purchase news from the monopolist.

In the absence of competition, the monopolist's problem collapses to a fairly straightforward dynamic programming problem. First, note that, since consumers can figure out the true value of d_t from n_t, the firm's choice of slant in equilibrium affects only its current period payoff and not the future state. Thus, the firm will choose the slanting policy that allows it to charge the highest possible price that period – this is the same as the monopolist's solution in MS (2005). Specifically,

$$s_t^*(d_t) = \gamma(b_t - d_t) \tag{3.9}$$

$$P_t^* = \bar{u} - \chi\gamma(b_t^2 + v_d) \tag{3.10}$$

Notice, importantly, that the price the firm can charge in equilibrium is strictly decreasing in the absolute value of b_t and strictly decreasing in v_t. This implies that the firm has at least two dynamic incentives:

(1) The firm prefers consumers to have as small a bias as possible and might be willing to forgo some profit today to keep future beliefs sufficiently close to zero.

[1] The case of a monopolist news firm with heterogeneous consumers is a trivial extension that mirrors the analogous case in MS (2005) without affecting the qualitative dynamics in which we are most interested.

(2) The firm prefers belief variance to be as low as possible. This gives the firm additional incentive to report news, since each news report decreases the variance of future beliefs.

Since we already know the slanting and pricing policies of the firm in equilibrium and that consumers will indeed purchase news, the monopolist's problem reduces to a dynamic programming problem with state vector (b_t, p_t). In each period, the firm merely observes d_t and decides whether to report news. Let $V(b_t, p_t)$ be the firm's value function and $V(b_t, p_t)|_{d_t}$ be the value function upon observing d_t. Then,

(3.11)
$$V(b_t, p_t)|_{d_t} = max\left\{\bar{u} - \chi\gamma\left(b_t^2 + \frac{1}{p_t} + \frac{1}{p_\epsilon}\right) + \delta V\left(\frac{p_t b_t + p_\epsilon d_t}{p_t + p_\epsilon}, p_t + p_\epsilon\right), \delta V(b_t, p_t)\right\}$$

(3.12)
$$V(b_t, p_t) = E_{d_t}[V(b_t, p_t)|_{d_t}]$$

Our analysis focuses on two issues: the nature of the firm's disclosure set (that is, how the news firm systematically censors data) and the evolution of consumer beliefs over time.

Lemma 3.1. *Holding p_t constant, $V(b_t, p_t)$ is single-peaked at $b_t = 0$ if D_t is non-empty.*

Lemma 3.2. *If D_t is non-empty, then it is convex and symmetric around the point $d_t^* \equiv -\frac{p_t b_t}{p_\epsilon}$.*

Lemma 3.1 tells us that the firm's long term expected payoff is highest when consumers are unbiased. This, in turn, tells us something about which data point is most desirable in the form of Lemma 3.2: the d_t that results in $b_{t+1} = 0$, which we denote by d_t^*. In particular, the closer d_t is to d_t^*, the closer b_{t+1} will be to zero. Since D_t will always be an interval symmetric around a point with a sign opposite that of b_t, if follows that this information disclosure policy will result in belief convergence to zero.

Proposition 3.1. *Consider a sequence of data $\{d_\tau\}_{\tau=t}^{\infty}$ such that D_τ is non-empty for every τ. Then, $E[b_\tau] \to 0$ and $D_\tau \to R$ as $\tau \to \infty$.*

Proposition 3.1 states that, as long as a monopolist continues to report news, bias will not persist. Moreover, in the long run, the monopolist will disclose all data to consumers.

3.3.2. Monopoly: Computational Analysis

In order to further investigate a monopolist's revelation incentives, we numerically solve for the firm's value function $V(b_t, p_t)$ and equilibrium disclosure strategy $D(b_t, p_t)$. Because the arguments require a bounded state space, however, we use state variables (b_t, v_t) instead of (b_t, p_t). While the precision p_t becomes unbounded with revelation, at any time $t + k \geq t$ we have $v_{t+k} \in [0, v_t]$. Additionally, the following lemma bounds the state variable b_t.

Lemma 3.3. *For all sequences $\{b_t\}_{t=1}^{\infty}$ for which D_t is non-empty for all t, b_t is bounded from above and below in both the monopolist and duopolist cases. Specifically:*

$$b_t \in \left(-\sqrt{\frac{\bar{u}}{\chi\gamma(1-\delta)}}, \sqrt{\frac{\bar{u}}{\chi\gamma(1-\delta)}}\right) \tag{3.13}$$

Given the bounded state space, standard arguments show that the operator

$$Tf(b_t, v_t) = E_{d_t}\left[\max\left\{\bar{u} - \chi\gamma\left(b_t^2 + v_t + v_\epsilon\right) + \delta f\left(\frac{b_t v_\epsilon + d v_t}{v_\epsilon + v_t}, \frac{v_t v_\epsilon}{v_t + v_\epsilon}\right), \delta f(b_t, v_t)\right\}\right]$$

is a contraction mapping [Blackwell (1965)] under the sup-norm. As a result, there exists a unique function V such that $TV = V$. Furthermore, this function can be computed by iteratively applying the operator T:

$$Tf^i(b_t, v_t) = f^{i-1}(b_t, v_t), \quad f^0(b_t, v_t) = f^0$$

until an arbitrary degree of accuracy ϵ is reached:

$$d(f^i, f^{i-1}) < \epsilon.$$

The numerical implementation proceeds by approximating V on a finite grid and by considering optimal revelation strategies on a discretized space of news. Specific details are relegated to Appendix D.

We now present numerical results for the parameter values $\phi = 0.7, \chi = 0.7, \delta = 0.95, \bar{u} = 5, v_\omega = 0.25, v_\epsilon = 0.25$. Figure F.1 gives the estimated value function. Figure F.2 gives the probability of revealing news d_t. These highlight the monopolist's incentives to disclose news in a way that creates an unbiased and well-informed (high precision) set of readers. In particular, given such a population of consumers, the firm reveals all news.

The next set of results concern convergence. Figures F.3 and F.4 show the convergence of v_t and b_t at a point $(b_t, v_t) = (10.75, 0.1)$ with a low probability of reporting and a low continuation value. In Figure F.3, the average bias and variance are plotted against the period number, whereas in Figure F.4, the average states are plotted against one another; the large black dots in Figure F.4 mark the average (b_t, v_t) pair at every tenth iteration. This gives an indication of the speed of convergence. For a low value point, convergence of the average path is slow. However, this is because on many paths, the probability of reporting remains near zero. Once news is revealed, the firm quickly begins to report news with probability near one. Either a point is not moving between periods, or the path quickly enters into a region where nearly all news is revealed. This also explains the observed near-linearity on the (b_t, v_t) plot in Figure F.4; a single path either converges quickly to $(0, 0)$ or remains at the original value until a sufficiently "shocking" piece of news is revealed.

In contrast, at a moderate value point such as $(b_t, v_t) = (8, 0.22)$, we see a much faster rate of convergence (Figures F.5 and F.6). Few paths do not report news in an early iteration, and once they begin reporting they quickly move toward $(0, 0)$. In Figure F.6 we also see that by the tenth iteration, the average path has shifted significantly toward the origin, and it only accelerates in later iterations.

3.4. Duopolist's Problem

3.4.1. Duopoly: Theory

Suppose there are up to two types of consumers who have current biases $b_1 \leq b_2$. Suppose further that two firms have already positioned themselves at locations $z_1 \leq z_2$, where "locating" implies $s_i = s^{z_i}(d)$ and that $d_t \in D_{1,t} \cap D_{2,t}$. [2] Then, conditional on P_j, firm i captures both types of consumers if and only if:

$$\bar{u} - \chi\gamma(b_j^2 + v_d) - \phi\gamma(z_i - b_j)^2 - P_i > \bar{u} - \chi\gamma(b_j^2 + v_d) - \phi\gamma(z_j - b_j)^2 - P_j \tag{3.14}$$

$$\iff P_i < P_j + \phi\gamma[(z_j + z_i)(z_j - z_i) + 2b_j(z_i - z_j)] \tag{3.15}$$

Specifically, for firm 1 to capture type 2 consumers, it must be that:

$$P_1 < P_2 - 2\phi\gamma\Delta z[b_2 - \bar{z}] \tag{3.16}$$

And for for firm 2 to capture type 1 consumers, it must be that:

$$P_2 < P_1 - 2\phi\gamma\Delta z[\bar{z} - b_1] \tag{3.17}$$

Where:

[2] As in MS (2005), we simply impose as an assumption that duopolists use linear slanting strategies.

$$\Delta z \equiv z_2 - z_1 \tag{3.18}$$

$$\bar{z} \equiv \frac{z_1 + z_2}{2} \tag{3.19}$$

Alternatively, if d_t is in the disclosure set of only one firm, the active firm captures the whole market only if:

$$P_1 \leq \bar{u} - \chi\gamma(b_2^2 + v_d) - \phi\gamma(z_1 - b_2)^2 \tag{3.20}$$

$$P_2 \leq \bar{u} - \chi\gamma(b_1^2 + v_d) - \phi\gamma(z_2 - b_1)^2 \tag{3.21}$$

These expressions capture some important differences between the monopolist case and the duopolist case. Firm j would ideally like to cater to the preferences of its half of the market by setting $z_j = b_j$ and charging the monopolist price to capture all of the surplus. However, if b_1 and b_2 are sufficiently close to each other, this creates an incentive for firm i to undercut this price and capture the entire market for itself. However, the case where one firm serves the entire market is never an equilibrium:

Lemma 3.4. *If $b_1 \neq b_2$, then each firm serves at most one type of consumer in equilibrium.*

Conditional on disclosure, the goal of each news firm in the stage game is to charge the highest possible price subject to two constraints. First, they have to charge a price low enough that consumers are willing to buy news. We will denote this as the WTP constraint:

$$P_1 \leq \bar{u} - \chi\gamma(b_1^2 + v_{d,1}) - \phi\gamma(z_1 - b_1)^2 \tag{3.22}$$

$$P_2 \leq \bar{u} - \chi\gamma(b_2^2 + v_{d,2}) - \phi\gamma(z_2 - b_2)^2 \tag{3.23}$$

In the monopolist game, this is the only constraint and therefore always binds; the monopolist always captures all the surplus when consumers are homogeneous. We will denote by P_i^m the price the monopolist would charge in state (b_i, v_i) if it were positioned at location z_i, and this price is equal to the right hand side of the WTP constraint for each respective firm.

The duopolist firm has a second constraint: conditional on z_j the price for firm i must be low enough that it is not profitable for firm j to undercut and steal the whole market. We will call this the undercut constraint:

$$P_1 \leq \frac{1}{2}P_2 + 2\phi\gamma\Delta z(\bar{z} - b_1) \tag{3.24}$$

$$P_2 \leq \frac{1}{2}P_1 + 2\phi\gamma\Delta z(b_2 - \bar{z}) \tag{3.25}$$

Some stage game examples will help demonstrate the dynamic tension between these two types of constraints. For these examples, assume the following:

$$\bar{u} = 10 \tag{3.26}$$

$$\phi = \chi = 2 \tag{3.27}$$

$$v_{d,1} = v_{d,2} = 1 \tag{3.28}$$

Example 3.1. $-\mathbf{b_1} = \mathbf{b_2} = \mathbf{2}$ *We know there will not be a stage game equilibrium where one firm captures the entire market. Let us begin by setting $z_i = b_i$, binding the WTP constraint for each firm, and then checking the respective undercut constraint. When $-z_1 = z_2 = 2$, we have $P_1^m = P_2^m = 5$. The undercut constraint for firm 1 is satisfied:*

$$5 \leq 2.5 + (2)(1)(4)(2) = 18.5 \tag{3.29}$$

And the undercut constraint for firm 2 is analogous. So we have a stage game equilibrium; neither firm can benefit by raising price (because consumers are not willing to pay more), lowering price (undercutting the other firm is too costly to be profitable), or by changing position (choosing any $z_i \neq b_i$ makes P_i^m smaller).

What does this tell us about dynamic incentives with respect to each firm's disclosure policy? The crux of this equilibrium is that the two consumer types are sufficiently differentiated that each firm operates as a de facto monopolist for its respective market segment. The local dynamic incentive here is the same as in the monopoly case: each firm could charge a higher price if its consumers beliefs were a little bit closer to zero. In fact, there is so much slack in the undercut constraint that, holding b_j constant, firm i

would be best off if $b_i = 0$. In the case where $b_i = 0$ and $b_j = 2$, $P_i^m = 9$ is still satisfies the undercut constraint by a substantial margin.

Of course, this local incentive to shift consumers closer to zero is not a global incentive. To see this, let us examine an extreme scenario:

Example 3.2. $\mathbf{b_1 = b_2 = 0}$ *In this case, consumers are homogeneous, and we will observe* $z_1 = z_2 = p_1 = p_2 = 0$. *This result follows from standard undifferentiated Bertrand reasoning and is shown explicitly in Proposition 3 of MS (2005). Because stage game profits are zero, the undercut constraint dominates the WTP constraint and the firms have a strong incentive to manipulate consumers apart from each other. For example, we could have disclosure sets such that:*

$$D_1 = \{d \in R : 0 \geq d \geq -\bar{d}\} \tag{3.30}$$

$$D_2 = \{d \in R : \bar{d} \geq d \geq 0\} \tag{3.31}$$

In this case, firm 1 only reports data points that lie left of center (but still above some extreme lower bound) and firm 2 only reports data points that lie right of center (but still below some extreme upper bound). For any given d_t such that $\bar{d} \geq |d_t| \neq 0$, half of the consumers will consume news and update to $b_{t+1} \neq 0$ while the other half will remain at $b_{t+1} = 0$. In this way, forward-looking, profit-maximizing news firms can actually create polarization and political conflict even where none previously existed. This analysis has a key implication: the worst possible scenario is for consumer types to have very similar beliefs with very low variance, so firms will choose disclosure policies that prevent this from happening. Thus, we will **not** observe all consumer beliefs converging to unbiasedness as we did in the monopoly case. It is plausible, however, that one type converges to unbiasedness while another converges to beliefs away from zero (more information on this scenario in the computational analysis below).

Example 3.3. $-\mathbf{b_1} = \mathbf{b_2} = \mathbf{1}$ *Lastly, let us consider this intermediate example. In this case,* $P_1^m = P_2^m = 8$*, and both sets of constraints bind simultaneously:*

$$8 \leq 4 + (2)(1)(2)(1) = 8 \tag{3.32}$$

Firm i does not want $|b_i|$ to increase – this would tighten the WTP constraint and reduce its price. Firm i also does not want $|b_i|$ to decrease – this would tighten the undercut constraint and reduce its price. Thus, holding b_j constant, firm i wants to keep its consumers' beliefs exactly as they are. This gives us a clue about what types of convergence we may observe in our computational analysis below.

3.4.2. Duopoly: Computational Analysis

We next present the algorithm used to compute the value function for the duopoly case. As previously, specific details of the implementation are relegated to Appendix D. For simplicity of notation, let $x_t^i = (b_t^i, v_t^i)$ denote the pair of state variables for firm i at time t, let x_t without firm superscript denote the pair $x_t = (x_t^1, x_t^2)$, and let $\bar{x}_t$ denote the flipped pair $\bar{x}_t = (x_t^2, x_t^1)$. Let $\rho(x^i, d, r^i)$ denote the state-update function which returns the updated beliefs and variance given by expression (D.1) in Appendix D when $r^i = 1$ (news is revealed) and which returns the identity x^i when $r^i = 0$ (news is not revealed). In addition to choosing prices according to the Duopoly Algorithm presented in Appendix E, firms choose whether to reveal $r_t^i(d,x) \in \{0,1\}$ each possible piece of news $d \in R$. Given the value function V^1 and Firm 2's revelation strategy $r_t^2(\cdot,x)$, Firm 1's equilibrium revelation strategy solves:

$$\forall d, r_t^1(d,x_t) \in \arg\max\{P^1(x_t) + \delta V^1(\rho(x_t^1,d,1),\rho(x_t^2,d,r_t^2(d,x_t)),$$
$$\delta V^1(x_t^1,\rho(x_t^2,d,r_t^2(d,x_t)))\} \tag{3.33}$$

A Nash equilibrium of the revelation game jointly solves (3.33) and the analogous condition for Firm 2. Note that value functions are subscripted by player for now. An important point is that the duopoly game will generally have a large multiplicity of disclosure equilibria, and a value function for each player is specific to the equilibrium selected. Note also that we subscript the player strategies r_t^i to allow for the possibility that firms choose different equilibria in different iterations for the same state variables. For asymmetric equilibria there is no basis for assuming these value functions will be equal.

Simultaneously solving for the revelation equilibrium at each iteration may be computationally intensive, and it additionally requires us to take a stance on equilibrium selection. Therefore we begin by searching for a symmetric disclosure equilibrium in the manner of Pakes and McGuire (1994). Namely, we solve for the value function for a single firm, say Firm 1, and at each iteration t we assume that Firm 2 plays as Firm 1 would have played at the same state variables in the previous round, and then we compute Firm 1's optimal best response:

$$\forall d, r_t(d, x_t) \in \arg\max\{P^1(x_t) + \delta V_{t-1}(\rho(x_t^1, d, 1), \rho(x_t^2, d, r_{t-1}(d, \bar{x}_t))),$$

$$\delta V_{t-1}(x_t^1, \rho(x_t^2, d, r_{t-1}(d, \bar{x}_t)))\} \tag{3.34}$$

Note that, if found, a solution $V = V^t = V^{t-1}$ and strategy profile $r = r_t = r_{t-1}$[3] constitutes a symmetric equilibrium of the original game where $r^1(d, x) = r(d, x)$ and $r^2(d, x) = r(d, \bar{x})$.

In this game, however, it becomes clear ex post that a pure strategy equilibrium cannot be symmetric. In particular, the symmetry cannot be expected to hold for $b_1 = b_2 \neq 0, v_1 = v_2$. Both firms have an incentive to move toward zero, but they also have an incentive to keep their clientele sufficiently separated to extract high surplus. For a given d_t, we should then expect three types of outcomes. The first two are that either both firms report news or that neither reports news. In such equilibria both firms have a dominant strategy to either report or not. These involve symmetric actions. The third, however, involves the pair of equilibria where exactly one news firm reports. Given that the competing firm reports, it is optimal for the other firm to avoid reporting, lest the two firms move too close in bias-variance space. We conjecture that each firm strictly prefers the equilibrium in which it reports news.

Nevertheless, for now this computation should provide a good first-order approximation to equilibrium and to the corresponding value functions.[4] We use the approximate

[3] Numerically speaking, a pair of arbitrarily close approximations.

[4] Intuitively, we are computing a sort of level-k reasoning outcome, where Firm 1 best responds to a firm that would take the same action.

value function to simulate trajectories of bias, which are computed by selecting a true equilibrium given the value function (which for now is generated in a "near" equilibrium).

We expect this to provide a good approximation for two reasons. The first is that, when consumers have the same bias, the current period profit is zero for each firm regardless of the revelation decision, so the incentives for firms preferring one outcome over the other involve only future payoffs. Second, this still involves a sensible best response to a sensible action and may approximate more of a mixed strategy equilibrium. Consider again a d_t in a given state such that Firm 1 reports if and only if Firm 2 does not. In the current computation, in one iteration over the value function, Firm 1 reports. Responding to this action in the next iteration, it does not. Iterating over this logic, what we capture is a sort of mixed strategy response where the firms mix over the equilibria.[5] It is clear, however, that this computation is not a contraction mapping because the strategies will eventually cycle as described, and correspondingly the value function will also vary when we are in a revelation iteration versus a non-revelation iteration.

3.4.3. Results

As described in the previous section, the current algorithm does not compute a formal equilibrium of the game. However, we conjecture and expect that the properties highlighted in this section are robust to the approximation, and that we will see the same phenomena in a true equilibrium. Figures F.7 and F.8 graph the value function for one firm as we respectively shift the bias and variance of the other firm's consumers. An interesting phenomenon is that a lower variance firm can defend an optimal point and

[5] But correspondence to an actual mixed strategy equilibrium is broken by the presence of a discount factor $\delta \neq 1$.

force the other firm to move its consumers to a more extreme point, but if both firms have sufficiently certain (low variance) consumers then they may be stuck in a trap where both firms would like to move away from one another but are no longer able to provide sufficiently extreme news to do so. This "bias trap" is evidenced in the lower right quadrant of Figure F.8.

In Figures F.9 and F.10, we analyze the following question: given the approximate value function, how do firms optimally behave in the future? In order to do so, we have to take a stance on equilibrium selection. In particular, we plot the equilibrium where the same firm always reveals (does not reveal) in the case of multiple stage game revelation equilibria. We expect that this will provide an upper (lower) envelope to the value functions for the two players, and therefore it is an interesting case to examine. Note that in both the case where firms start off at the same b_i and the case where they start off equidistantly on opposing sides of 0, the consumers of the firm with the reporting advantage eventually have correct beliefs. The other firm quickly stops revealing news that would move it too close to this firm in equilibrium. In a sense the firm with a first-mover advantage ("first to press") can move beliefs to the truth and force the other firm to reveal only news that maintains bias. Whether the firms remain separated in all equilibria for all discount factors δ is not clear from this figure, although the figures seem to indicate a strong and persistent incentive for the firms to remain differentiated – namely, a firm will stop revealing news if doing so moves it too close to the other firm. Another interesting question is: does one firm necessarily move to the truth, i.e. do there exist equilibria without convergence to zero for at least one firm as $t \to \infty$? We think in general the answer to the latter question will be yes, and this depends on the equilibrium we select.

3.5. Conclusion

In this essay, we show how competitive forces in the news media industry can create political division and polarization among consumers, even in cases where none initially existed. For policy makers, this is a negative outcome in two ways. First, it results in decreased welfare within the confines of the model. In our framework, total surplus is maximized when when consumers are homogeneous and unbiased. A monopolist gets us to this outcome, though the monopolist also reaps all of the surplus. When duopolists shift consumers away from unbiasedness as they do in our model, they are decreasing total surplus in order to increase producer surplus. Second, there are obvious, unmodeled negative externalities associated with political conflict that are not internalized by news firms (if anything, such conflict may increase demand for news media).

Unfortunately, highlighting this problem does not lead to any easy solution. Although the monopolist case seems to lead to a positive outcome, this result depends crucially on the assumption that the monopolist is motivated solely by profit. Democracies value competitive news industries specifically because this assumption may not always be true. One possible approach that could improve outcomes is to use education to create consumers with low-bias, high-confidence initial beliefs, but such things are easier said than done.

There are natural avenues to continue this line of research. Within the context of theory and computation, we can continue to consider new cases. There is room for variation along several dimensions, including within the parameter space, the set of initial states, and the set of equilibrium selection assumptions. Robust case analysis can help us better understand when the issues raised in this paper are most severe.

Moreover, virtually all of the theoretical predictions in this model are testable empirically. Media scholars have means to measure political slant, and the methodology of Gentzkow and Shapiro (2010) has grown quite popular. Scholars also have access to rich data about consumer beliefs through organizations such as the Pew Research Center. It is our hope that continued research will paint a clearer picture about the relationship between news media and consumer beliefs and shed some light on how to mitigate the concerns raised in this essay.

References

[1] Alcalde, J. (1996). Implementation of Stable Solutions to Marriage Problems. *Journal of Economic Theory*, 69(1), 240-254.

[2] Altonji, J. G., and Pierret, C. R. (2001). Employer Learning and Statistical Discrimination. *Quarterly Journal of Economics*, 116(1), 313-350.

[3] Ambrus, A., and Argenziano, R. (2009). Asymmetric Networks in Two-Sided Markets. *American Economic Journal: Microeconomics*, 1(1), 17-52.

[4] Anderson, S., and McLaren J. (2012). Media Mergers and Media Bias with Rational Consumers. *Journal of the European Economic Association*, 10(4), 831-859.

[5] Anwar, S., and Fang, H. (2006). An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence. *American Economic Review*, 96(1), 127-151.

[6] Antonovics, K., and Knight, B. G. (2009). A New Look at Racial Profiling: Evidence from the Boston Police Department. *Review of Economics and Statistics*, 91(1), 163-177.

[7] Armstrong, M. (2006). Competition in Two-Sided Markets. *The RAND Journal of Economics*, 37(3), 668-691.

[8] Arnosti, N. (2016). Centralized Clearinghouse Design: A Quantity-Quality Tradeoff. http://dx.doi.org/10.2139/ssrn.2571527

[9] Arrow, K. J. (1973). The Theory of Discrimination. In Ashenfelter, O. and Rees, A. (Eds.) *Discrimination in Labor Markets*. Princeton University Press.

[10] Ashlagi, I., Kanoria, Y., and Leshno, J. D. (2017). Unbalanced Random Matching Markets: The Stark Effect of Competition. *Journal of Political Economy*, 125(1), 69-98.

[11] Athey, S., and Levin, J. (2017). The Value of Information in Monotone Decision Problems. *Research in Economics*. doi: 10.1016/j.rie.2017.01.001.

[12] Ayres, I., and Waldfogel, J. (1994). A Market Test for Race Discrimination in Bail Setting. *Stanford Law Review*, 46(5), 987-1047.

[13] Azevedo, E. M., and Budish, E. (2017). Strategy-Proofness in the Large. NBER working paper no. 23771.

[14] Baron, D. P. (2006). Persistent Media Bias. *Journal of Public Economics*, 90(1), 1-36.

[15] Becker, G. S. (1957). *The Economics of Discrimination*. University of Chicago Press.

[16] Bernhardt, D., Krasa, S., and Polborn, M. (2008). Political Polarization and the Electoral Effects of Media Bias. *Journal of Public Economics*, 92(5), 1092-1104.

[17] Bischoff, A., Bovier, P. A., Isah, R., Franoise, G., Ariel, E., and Louis, L. (2003). Language Barriers Between Nurses and Asylum Seekers: Their Impact on Symptom Reporting and Referral. *Social Science and Medicine*, 57(3), 503-512.

[18] Blackwell, D. (1951). Comparison of Experiments. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. The Regents of the University of California.

[19] Blackwell, D. (1965). Discounted Dynamic Programming. *Annals of Mathematical Statistics*, 36(1), 226-235.

[20] Blomqvist, N. (1952). On an Exhaustion Process. *Scandinavian Actuarial Journal*, 1952(3-4), 201-210.

[21] Brigham, J. C., and Barkowitz, P. (1978). Do "They All Look Alike?" The Effect of Race, Sex, Experience, and Attitudes on the Ability to Recognize Faces. *Journal of Applied Social Psychology*, 8(4), 306-318.

[22] Brown, A. L., Camerer, C. F., and Lovallo, D. (2012). To Review or Not to Review? Limited Strategic Thinking at the Movie Box Office. *American Economic Journal: Microeconomics*, 4(2), 1-26.

[23] Burke, J. (2008). Primetime Spin: Media Bias and Belief Confirming Information. *Journal of Economics and Management Strategy*, 17(3), 633-665.

[24] Cacioppo, J. T., Cacioppo, S., Gonzaga, G. C., Ogburn, E. L., and VanderWeele, T. J. (2013). Marital Satisfaction and Break-Ups Differ Across On-Line and Off-Line Meeting Venues. *Proceedings of the National Academy of Sciences*, 110(25), 10135-10140.

[25] Caillaud, B., and Jullien, B. (2003). Chicken and Egg: Competition Among Intermediation Service Providers. *RAND Journal of Economics*, 309-328.

[26] Chan, J., and Suen, W. (2008). A Spatial Theory of News Consumption and Electoral Competition. *Review of Economic Studies*, 75(3), 699-728.

[27] Chan, J., and Suen W. (2009). Media as Watchdogs: The Role of News Media in Electoral Competition. *European Economic Review*, 53, 799-814.

[28] Charles, K. K., and Guryan, J. (2011). Studying Discrimination: Fundamental Challenges and Recent Progress. *Annual Review of Economics*, 3(1), 479-511.

[29] Coate, S., and Loury, G. C. (1993). Will Affirmative-Action Policies Eliminate Negative Stereotypes?. *American Economic Review*, 83(5), 1220-1240.

[30] Damiano, E., and Li, H. (2007). Price Discrimination and Efficient Matching. *Economic Theory*, 30(2), 243-263.

[31] Damiano, E., and Hao, L. (2008). Competing Matchmaking. *Journal of the European Economic Association*, 6(4), 789-818.

[32] DellaVigna, S., and Kaplan, E. (2007). The Fox News Effect: Media Bias and Voting. *Quarterly Journal of Economics*, 122(3), 1187-1234.

[33] Dominitz, J., and Knowles, J. (2006). Crime Minimisation and Racial Bias: What Can We Learn from Police Search Data?. *Economic Journal*, 116(515), 368-384.

[34] Donohue III, J. J., and Levitt, S. D. (2001). The Impact of Race on Policing and Arrests. *Journal of Law and Economics*, 44(2), 367-394.

[35] Druckman, J. N., and Parkin, M. (2005). The Impact of Media Bias: How Editorial Slant Affects Voters. *Journal of Politics*, 67(4), 1030-1049.

[36] Duggan, J., and Martinelli C. (2011). A Spatial Theory of Media Slant and Voter Choice. *Review of Economic Studies*, 78, 640-666.

[37] Ebrahimy, E., and Shimer, R. (2010). Stock-Flow Matching. *Journal of Economic Theory*, 145(4), 1325-1353.

[38] Erdös, P., and Renyi, A. (1964). On Random Matrices. *Magyar Tud. Akad. Mat. Kutató Int. Közl*, 8(455-461).

[39] Gale, D., and Shapley, L. S. (1962). College Admissions and the Stability of Marriage. *American Mathematical Monthly*, 69(1), 9-15.

[40] Gale, D., and Sotomayor, M. (1985a). Some Remarks on the Stable Matching Problem. *Discrete Applied Mathematics*, 11(3), 223-232.

[41] Gale, D., and Sotomayor, M. (1985b). Ms. Machiavelli and the Stable Matching Problem. *American Mathematical Monthly*, 92(4), 261-268.

[42] Gentzkow, M., and Shapiro, J. M. (2006). Media Bias and Reputation. *Journal of Political Economy*, 114(2), 280-316.

[43] Gentzkow, M., and Shapiro J. M. (2010). What Drives Media Slant? Evidence from U.S. Daily Newspapers. *Econometrica*, 78(1), 35-71.

[44] Gentzkow, M., Shapiro J.M., and Stone D. (2015). Media Bias in the Marketplace: Theory. In Anderson S., Waldfogel J., and Stromberg D. (Eds.) *Handbook of Media Economics, Volume 1B*. Elsevier B.V.

[45] Gerber, A. S., Karlan, D., and Bergan, D. (2009). Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions. *American Economic Journal: Applied Economics*, 1(2), 35-52.

[46] Groseclose, T., and Milyo, J. (2005). A Measure of Media Bias. *Quarterly Journal of Economics*, 120(4), 1191-1237.

[47] Hernández-Murillo, R., and Knowles, J. (2004). Racial Profiling or Racist Policing? Bounds Tests in Aggregate Data. *International Economic Review*, 45(3), 959-989.

[48] Hitsch, G.J., Hortasu, A., and Ariely, D. (2010). Matching and Sorting in Online Dating. *American Economic Review*, 100(1), 130-163.

[49] Hmielowski, J. D., Beam, M. A., and Hutchens, M. J. (2015). Structural Changes in Media and Attitude Polarization: Examining the Contributions of TV News Before and After the Telecommunications Act of 1996. *International Journal of Public Opinion Research*, 28(2), 153-172.

[50] Immorlica, N., and Mahdian, M. (2005, January). Marriage, Honesty, and Stability. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (pp. 53-62). Society for Industrial and Applied Mathematics.

[51] Jin, G. Z., Luca M., and Martin D. (2015). Is No News (Perceived As) Bad News? An Experimental Investigation of Information Disclosure. NBER working paper no. 21099.

[52] Judd, K. L. (1998). *Numerical Methods in Economics*. Cambridge, Mass.: MIT press.

[53] Kara, T., and Sönmez, T. (1996). Nash Implementation of Matching Rules. *Journal of Economic Theory*, 68(2), 425-439.

[54] Kemp, A. W. (1998). Absorption Sampling and the Absorption Distribution. *Journal of Applied Probability*, 35(2), 489-494.

[55] Knowles, J., Persico, N., and Todd, P. (2001). Racial Bias in Motor Vehicle Searches: Theory and Evidence. *Journal of Political Economy*, 109(1), 203-229.

[56] Kojima, F., and Pathak, P. A. (2009). Incentives and Stability in Large Two-Sided Matching Markets. *American Economic Review*, 99(3), 608-627.

[57] Lehmann, E. L. (1988). Comparing Location Experiments. *Annals of Statistics*, 16(2), 521-533.

[58] Malpass, R. S., and Kravitz, J. (1969). Recognition for Faces of Own and Other Race. *Journal of Personality and Social Psychology*, 13(4), 330-334.

[59] Manski, C. F. (2003). *Partial Identification of Probability Distributions*. Springer-Verlag.

[60] Imbens, G. W., and Manski, C. F. (2004). Confidence Intervals for Partially Identified Parameters. *Econometrica*, 72(6), 1845-1857.

[61] Manski, C. F. (2006). Search Profiling with Partial Knowledge of Deterrence. *Economic Journal*, 116(515), 385-401.

[62] Martin, G. J., and A. Yurukoglu (2017). Bias in Cable News: Persuasion and Polarization. *American Economic Review*, forthcoming.

[63] McVitie, D. G., and Wilson, L. B. (1970). Stable Marriage Assignment for Unequal Sets. *BIT Numerical Mathematics*, 10(3), 295-309.

[64] Mullainathan, S., and Shleifer, A. (2005). The Market for News. *American Economic Review*, 95(4), 1031-1053.

[65] National Research Council. (2004). *Measuring Racial Discrimination*. National Academies Press.

[66] Pakes, A., and McGuire P. (1994). Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model. *RAND Journal of Economics*, 25, 555-589.

[67] Persico, N., and Todd, P. (2006). Generalising the Hit Rates Test for Racial Bias in Law Enforcement, with an Application to Vehicle Searches in Wichita. *Economic Journal*, 116(515), 351-367.

[68] Phelps, E. S. (1972). The Statistical Theory of Racism and Sexism. *American Economic Review*, 62(4), 659-661.

[69] Rochet, J. C., and Tirole, J. (2003). Platform Competition in Two-Sided Markets. *Journal of the European Economic Association*, 1(4), 990-1029.

[70] Rochet, J. C., and Tirole, J. (2006). Two-Sided Markets: A Progress Report. *The RAND Journal of Economics*, 37(3), 645-667.

[71] Roth, A. E. (1984a). Misrepresentation and Stability in the Marriage Problem. *Journal of Economic Theory*, 34(2), 383-387.

[72] Roth, A. E. (1984b). The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory. *Journal of Political Economy*, 92(6), 991-1016.

[73] Roth, A. E. (1986). On the Allocation of Residents to Rural Hospitals: A General Property of Two-Sided Matching Markets. *Econometrica*, 425-427.

[74] Roth, A. E. (2002). The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics. *Econometrica*, 70(4), 1341-1378.

[75] Roth, A. E., and Peranson, E. (1999). The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design. *American Economic Review*, 89(4), 748.

[76] Stone, D. F. (2011). Ideological Media Bias. *Journal of Economic Behavior and Organization*, 78, 256-271.

[77] Strömberg, D. (2004). Mass Media Competition, Political Competition, and Public Policy. *Review of Economic Studies*, 71, 265-284.

[78] Tamer, E. (2010). Partial Identification in Econometrics. *Annual Review of Economics*, 2(1), 167-195.

[79] Weyl, E. G. (2010). A Price Theory of Multi-Sided Platforms. *American Economic Review*, 100(4), 1642-1672.

APPENDIX A

An Absolute Test for Racial Prejudice: Informativeness

Theorem A.1 (Athey and Levin (2017), Theorem 1)**.** *Suppose signal structures F and F' satisfy FOSD. Then F' is preferred to F by all treators with a utility function that is supermodular in (a, q) if and only if:*

$$F'_Q(\cdot|F'_X(X) > \alpha) \succ_{FOSD} F_Q(\cdot|F_X(X) > \alpha) \quad \textit{for all } \alpha \in [0, 1]. \tag{A.1}$$

Proof of Lemma 1.2. If $F^{r,t} = F^{r,t'}$, the result is trivial. Therefore suppose $F^{r,t} \succ F^{r,t'}$. Fix any average action $\bar{a}$. By definition of the information structure ordering (A.1),

$$F_Q^{r,t}(\cdot|F_X^{r,t}(X) \geq 1 - \bar{a}) \succ_{FOSD} F_Q^{r,t'}(\cdot|F_X^{r,t'}(X) \geq 1 - \bar{a}).^1$$

By FOSD and because the identity function is trivially increasing,

$$\mathbb{E}_Q^{r,t}[Q|F_X^{r,t}(X) \geq 1 - \bar{a})] \geq \mathbb{E}_Q^{r,t'}[Q|F_X^{r,t'}(X) \geq 1 - \bar{a})]. \tag{A.2}$$

Equivalently,

$$k^{r,t}(\bar{a}) \geq k^{r,t'}(\bar{a}).$$

Equivalence follows because a signal is included in the conditioning statements of the expectations in (A.2) if and only if the treator – applying an optimal decision rule that generates aggregate search rate $\bar{a}$ – would treat such a signal. $\square$

[1]The weak inequality is without loss since I assume the set of signal outcomes is atomless.

APPENDIX B

Revenue from Matching Platforms: Omitted Proofs

Proof of Theorem 2.1. To randomly generate a realization of K^{DA}, the process would randomly generate preferences, and then run DA. However this is probabilistically equivalent to the following process:

(i) randomly determine whether each w is acceptable to each m (given p_m),

(ii) randomly order each m's acceptable women to determine ordinal preferences, and then

(iii) run DA.

To randomly generate a realization of $K^{r\mathrm{MAP}}$, the process would randomly generate preferences and $\rhd_m$'s, and then run MAP. However this is probabilistically equivalent to the following process:

(i) randomly generate whether each w is acceptable to each m (given p_m),

(ii) randomly order each m's acceptable women to determine a relative meeting order over just those women,

(iii) run DA, using these meeting orders as "preferences".

This process skips the steps in which men meet unacceptable women. It is clear, however, that this would be a redundant step[1] hence the two processes are equivalent. □

Proof of Lemma 2.1. As in the matching literature, we can consider the equivalent case where men propose sequentially, and the current proposer is given by the lowest

[1]We intentionally include this redundancy since it allows derivation of our later results.

indexed man who is currently unengaged. We will define a recursive proposal probability function $\pi_k(\eta, \mathcal{P}, \mu)$ to compute the probability of k matches conditional on:

- A W-length vector η whose ith entry counts the number of compatible proposals received so far by woman i.
- An order $\mathcal{P}$ of meetings remaining for each of the men.[2]
- A temporary match μ which records the set of engagements at the given stage.[3]

Note that, given our lowest-unengaged-index proposal rule, the remaining meetings $\mathcal{P}$ and the temporary match μ are enough to determine the next proposer $m(\mathcal{P}, \mu)$, as well as the woman $w(\mathcal{P}, \mu)$ he meets, in the next step of a MAP algorithm with remaining meetings $\mathcal{P}$ and a temporary match μ. Additionally, let $\delta(\mathcal{P}, \mu)$ denote the transformation which deletes the current meeting between m and w from the meeting order, and let $\rho(\mathcal{P}, \mu)$ denote the transformation which updates the temporary match μ by replacing w's current engagement (if any) by $m(\mathcal{P}, \mu)$.

Then observe that we can write the conditional probability π_k of k matches recursively as follows:

$$\begin{aligned}\pi_k(\eta, \mathcal{P}, \mu) =& q\pi_k(\eta, \mathcal{P}', \mu) \\ &+ (1-q)\left(1 - \frac{1}{n_w'}\right)\pi_k(n', \mathcal{P}', \mu) \\ &+ (1-q)\frac{1}{n_w'}\pi_k(n', \mathcal{P}', \mu')\end{aligned}$$

[2]We use $\mathcal{P}$ because we can think without loss of this object lying in the space of (the men's) preferences. Namely, think of the meetings left for any man in the same way as a preference, and denote the end of the meeting list in the same way as an IR constraint.

[3] Again, this μ can inhabit the same space as a traditional matching function μ.

where $w = w(\mathcal{P}, \mu)$, $n' = n + \mathbb{1}_w$ denotes the new vector in which we have augmented the wth value of n by 1, $\mathcal{P}' = \delta(\mathcal{P}, \mu)$, and $\mu' = \rho(\mathcal{P}, \mu)$.

The three terms correspond to the three possible outcomes at such a stage of MAP given random preferences:

- m and w are incompatible, which occurs with probability q. To proceed we simply remove w from m's meeting order.
- m and w are compatible but w prefers her previous engagement. This occurs with probability $(1-q)\left(1 - \frac{1}{n'_w}\right)$. We update the meeting order and the number of compatible men that women w has met.
- m and w are compatible and w prefers m to her previous engagement. This occurs with probability $(1-q)\frac{1}{n'_w}$. We update the meeting order and the number of compatible men met, and we replace w's previous engagement with m in the temporary match.

To close the loop, we specify the terminal conditions:

$$\pi_k(\eta, \emptyset, \mu) = \begin{cases} 1 & \text{if } |\mu| = k \\ 0 & \text{else.} \end{cases}$$

Note that the number of remaining meetings decreases by one at every step, so that every terminal node of the recursive computation is reached in finite steps.

We finish the proof by observing that the probability of k matches in a random economy with meeting order $\rhd$ is given by:

$$P(k, \rhd) = \pi_k(\eta_0, \rhd, \mu_0)$$

where η_0 is a length-W zero vector and μ_0 denotes the null match where all agents are single. It follows that the distribution of the number of matches is parametrized by the single variable q. $\square$

Proof of Lemma 2.2. We prove the result for MAP, and the result for DA follows immediately. Fix a profile of meeting orders $\rhd$, a value of $q \in (0, 1]$, and a profile of preferences u. Consider prices $p_m = 0$, $p_w = q$, and $p'_w = q' > q$ (for which $q(p_m, p_w) = q$, $q(p_m, p'_w) = q'$). At $p_m = 0$, each man finds every woman acceptable. Hence we can reinterpret each $\rhd_m$ as an ordinal preference relation over W, and reinterpret MAP($\rhd$) as Deferred Acceptance.[4] The input to DA is (i) the men's preferences in the form of $\rhd$, and (ii) the women's preferences in the form of their rankings over acceptable men at prices p_w and p'_w respectively.

The women's ordinal preferences over men at price p'_w are a truncation of their preferences at p_m. It follows from Theorem 2 of Gale and Sotomayor (1985)[5] that any agent who is unmatched when DA on the women's preferences derived from p_w remains unmatched when running DA on the women's preferences derived from p'_w. This proves monotonicity of the number of marriages for any fixed u. Hence taking expectations over all preferences u, we have $\bar{K}(q) > \bar{K}(q')$. $\square$

Proof of Theorem 2.3. Equation 2.3 can be rewritten using Lemma 2.1 (or Theorem 2.2 for the case of DA), along with a transformation of variables

$$\max_{p_m, p_w} (p_m + p_w)\bar{K}(q(p_m, p_w)) = \max_{p, p_m} p\bar{K}(q(p_m, p - p_m))$$

[4]Each m meets women in order of $\rhd_m$, and bothers proposing only if the woman finds him acceptable; each woman keeps her best proposer.

[5]On the stable matching problem, discrete applied math.

where $p \equiv p_m + p_w$ is the revenue from a single marriage. Observe that if (p^*, p_m^*) solves the latter maximization problem, then p_m^* must solve

$$\max_{p_m} p^* \bar{K}(q(p_m, p^* - p_m))$$

That is for fixed p, p_m should be chosen merely to maximize the expected number of marriages $\bar{K}()$. Consider first the general case for distributions F_m, F_w. By Lemma 2.2, this means p_m should be chosen to minimize $q(p_m, p^* - p_m) = F_m(p_m) + F_w(p^* - p_m) - F_m(p_m)F_w(p^* - p_m)$.

Differentiating $q()$ we have

$$\begin{aligned}\frac{\partial q(p_m, p^* - p_m)}{\partial p_m} &= (1 - F_w(p^* - p_m))f_m(p_m) - (1 - F_m(p_m))f_w(p - p_m) \\ &= (1 - F_m(p_m))(1 - F_w(p^* - p_m))\left[\frac{f_m(p_m)}{1 - F_m(p_m)} - \frac{f_w(p^* - p_m)}{1 - F_w(p^* - p_m)}\right]\end{aligned}$$

Now consider the case $F_m = F_w = F$.[6] By the monotone hazard rate condition this implies

$$\frac{\partial q(p_m, p^* - p_m)}{\partial p_m} \begin{cases} \leq 0 & \text{for } p_m < \frac{p^*}{2} \\ = 0 & \text{for } p_m = \frac{p^*}{2} \\ \geq 0 & \text{for } p_m > \frac{p^*}{2}. \end{cases}$$

Therefore $q(p_m, p^* - p_m)$ is minimized at $p_m = p^*/2$. □

Proof of Theorem 2.4. Fix M and W, and a profile of identical meeting orders. we want to know the probability $P(k; M, W)$ that this procedure ends with k couples. Clearly

[6]Of course when $F_m \neq F_w$ the optimal prices may not be equal, but the same principle of equating hazard rates applies.

Equation 2.4 holds whenever $M = 1$: the lone man in the economy is either incompatible with each woman ($P(0;1,W) = q^W$) or is not ($P(1;1,W) = 1 - q^W$).

Using induction on the number of men M, suppose that for any k, Equation 2.4 accurately describes $P(k; M-1, W)$. By the construction of the MAP algorithm with identical orders, and symmetry of the men, we have the following observation: Fixing M, consider running the algorithm only until man m_{M-1} is matched (or is rejected by all women); call this *the end of stage* $M-1$. The probability that k of the first $M-1$ men are married *at this point in the algorithm* is precisely $P(k; M-1, W)$, since a complete run of the algorithm for a randomized economy of size $(M-1, W)$ is equivalent to a run of the algorithm to the end of stage $M-1$ for a randomized economy of size (M, W).

Furthermore for the economy (M, W) to end up with k marriages it must be that, at the end of stage $M-1$, there were either k or $k-1$ temporary marriages. We separately consider these two cases.

Case 1: at the end of stage $M-1$, k men are temporarily matched. There are thus $W-k$ women currently unmatched. The algorithm now introduces man m_M, who begins to sequentially meet women. If w_1 is currently unmatched, there is probability $(1-q)$ that she accepts a proposal from m_M (ending the algorithm), and probability q that he must continue by meeting w_2 (if she exists). But if w_1 was temporarily matched, then with certainty *some* man—either m_M or her temporary partner—will be permanently matched to her, and the other man continues by meeting w_2 (if she exists). In this latter case, it is probabilistically irrelevant *which* man continues on to meet w_2 (by the i.i.d. assumption on utilities).

This process continues for each woman in turn until the algorithm ends. Each temporarily married woman w_j keeps some offer and sends the other man on to meet w_{j+1}. Each currently single woman (if met) ends the algorithm with an accepted proposal with probability $(1-q)$. Therefore, "stage M" does *not* add an additional marriage to the already existing k marriages with probability q^{W-k}.

Case 2: at the end of stage $M-1$, $k-1$ men are temporarily matched. There are thus $W-k+1$ women currently unmatched. As above, the introduction of man m_M in stage M fails to yield an additional match precisely when each of the $W-k+1$ is incompatible with the unique man who proposes to her. Therefore, "stage M" adds an additional marriage to the already existing $k-1$ marriages with probability $1-q^{W-k+1}$.

Combining Case 2 and Case 1 respectively, $P(k;M,W)$ equals

$$P(k-1;M-1,W)\cdot(1-q^{W-k+1})+P(k;M-1,W)\cdot q^{W-k}$$

Using Equation 2.4 to substitute for $P(\,\cdot\,;M-1,W)$ this becomes

$$
\begin{aligned}
&(1-q)^{k-1}q^{(M-k)(W-k+1)}\begin{bmatrix}M-1\\k-1\end{bmatrix}_q\begin{bmatrix}W\\k-1\end{bmatrix}_q[k-1]_q!(1-q^{W-k+1})\\
&\quad+(1-q)^k q^{(M-k-1)(W-k)}\begin{bmatrix}M-1\\k\end{bmatrix}_q\begin{bmatrix}W\\k\end{bmatrix}_q[k]_q!(q^{W-k})\\
=&\left(\frac{q^{M-k}}{[k]_q(1-q)}\right)(1-q)^k q^{(M-k)(W-k)}\begin{bmatrix}M-1\\k-1\end{bmatrix}_q\begin{bmatrix}W\\k-1\end{bmatrix}_q[k]_q!(1-q^{W-k+1})\\
&\quad+(1-q)^k q^{(M-k)(W-k)}\begin{bmatrix}M-1\\k\end{bmatrix}_q\begin{bmatrix}W\\k\end{bmatrix}_q[k]_q!\\
=&\left(\frac{q^{M-k}}{[k]_q(1-q)}\right)(1-q)^k q^{(M-k)(W-k)}\begin{bmatrix}M\\k\end{bmatrix}_q\frac{[k]_q}{[M]_q}\begin{bmatrix}W\\k\end{bmatrix}_q\frac{[k]_q}{[W-k+1]_q}[k]_q!(1-q^{W-k+1})\\
&\quad+(1-q)^k q^{(M-k)(W-k)}\begin{bmatrix}M\\k\end{bmatrix}_q\frac{[M-k]_q}{[M]_q}\begin{bmatrix}W\\k\end{bmatrix}_q[k]_q!\\
=&\,q^{M-k}(1-q)^k q^{(M-k)(W-k)}\begin{bmatrix}M\\k\end{bmatrix}_q\frac{[k]_q}{[M]_q}\begin{bmatrix}W\\k\end{bmatrix}_q[k]_q!\\
&\quad+(1-q)^k q^{(M-k)(W-k)}\begin{bmatrix}M\\k\end{bmatrix}_q\frac{[M-k]_q}{[M]_q}\begin{bmatrix}W\\k\end{bmatrix}_q[k]_q!\\
=&\,(1-q)^k q^{(M-k)(W-k)}\begin{bmatrix}M\\k\end{bmatrix}_q\begin{bmatrix}W\\k\end{bmatrix}_q[k]_q!\left(\frac{q^{M-k}[k]_q+[M-k]_q}{[M]_q}\right)\\
=&\,(1-q)^k q^{(M-k)(W-k)}\begin{bmatrix}M\\k\end{bmatrix}_q\begin{bmatrix}W\\k\end{bmatrix}_q[k]_q!\left(\frac{[M]_q}{[M]_q}\right) \qquad \square
\end{aligned}
$$

Proof of Equation 2.8. We rewrite Equation 2.4 in terms of k, $g \equiv M - k$, and $h \equiv W - k$, and take the limit as $k \to \infty$.

$$
\begin{aligned}
&\lim_{k\to\infty} P(k; k+g, k+h) \\
&= \lim_{k\to\infty} (1-q)^k q^{gh} \begin{bmatrix} k+g \\ k \end{bmatrix}_q \begin{bmatrix} k+h \\ k \end{bmatrix}_q [k]_q! \\
&= \lim_{k\to\infty} (1-q)^k q^{gh} \left(\frac{(1-q^{k+1}) \cdots (1-q^{k+g})}{(1-q)\cdots(1-q^g)} \right) \left(\frac{(1-q^{k+1}) \cdots (1-q^{k+h})}{(1-q)\cdots(1-q^h)} \right) [k]_q! \\
&= \lim_{k\to\infty} (1-q^k) \cdots (1-q) q^{gh} \left(\frac{(1-q^{k+1}) \cdots (1-q^{k+g})}{(1-q)\cdots(1-q^g)} \right) \left(\frac{(1-q^{k+1}) \cdots (1-q^{k+h})}{(1-q)\cdots(1-q^h)} \right) \\
&= \phi(q) q^{gh} \frac{1}{(1-q)\cdots(1-q^g)\cdot(1-q)\cdots(1-q^h)} \qquad \square
\end{aligned}
$$

Proof of Theorem 2.6. Follows from Equation 2.8. $\square$

APPENDIX C

A Dynamic Theory of Political Slant: Omitted Proofs

Proof of Lemma 3.1. $V(b_t, p_t)$, when fully expanded, is an infinite discounted stream of prices (or in an extreme scenario where D_t is empty, equal to 0). The price the firm is able to charge in any period $\hat{t} \geq t$ is strictly decreasing in $|b_{\hat{t}}|$. Moreover, $b_{\hat{t}}$ can be written as follows:

$$b_{\hat{t}} = \frac{p_t b_t + p_\epsilon \sum_{\tau=t}^{\hat{t}-1} d_\tau 1_{D_\tau}}{p_t + p_\epsilon \sum_{\tau=t}^{\hat{t}-1} 1_{D_\tau}}$$

Thus, $b_{\hat{t}}$ is strictly increasing in b_t and therefore prices in all periods after t are strictly decreasing in $|b_t|$. Hence, $V(b_t, p_t)$ is strictly decreasing in $|b_t|$ for all (b_t, p_t) such that D_t is non-empty.

□

Proof of Lemma 3.2. First, notice that P_t^* and $\delta V(b_t, p_t)$ are independent of d_t. It must be, then, that $V\left(\frac{p_t b_t + p_\epsilon d_t}{p_t + p_\epsilon}, p_t + p_\epsilon\right) \geq V\left(\frac{p_t b_t + p_\epsilon \tilde{d}_t}{p_t + p_\epsilon}, p_t + p_\epsilon\right)$ for all $d_t \in D_t$ and all $\tilde{d}_t \notin D_t$. We know from Lemma 3.1 that V is maximized at $b = 0$. Thus, since $p_t b_t + p_\epsilon d_t^* = 0$, it must be that $d_t^* \in D_t$.

To show that D_t is convex, consider some point $d_t \in D_t$. Let $\hat{d}_t = \alpha d_t + (1-\alpha) d_t^*$ be a convex combination of d_t and d_t^*. Then, $\left|\frac{p_t b_t + p_\epsilon \hat{d}_t}{p_t + p_\epsilon}\right| < \left|\frac{p_t b_t + p_\epsilon d_t}{p_t + p_\epsilon}\right|$, which implies $V\left(\frac{p_t b_t + p_\epsilon \hat{d}_t}{p_t + p_\epsilon}, p_t + p_\epsilon\right) \geq V\left(\frac{p_t b_t + p_\epsilon d_t}{p_t + p_\epsilon}, p_t + p_\epsilon\right)$. Therefore, it must be that $\hat{d}_t \in D_t$.

To show that D_t is symmetric around d_t^*, consider some point $d_t \in D_t$ such that $d_t = d_t^* + \alpha$, where α is some real number. Let $\hat{d}_t = d_t - \alpha$. Then, $\left|\frac{p_t b_t + p_\epsilon \hat{d}_t}{p_t + p_\epsilon}\right| = \left|\frac{p_t b_t + p_\epsilon d_t}{p_t + p_\epsilon}\right|$, which implies $V\left(\frac{p_t b_t + p_\epsilon \hat{d}_t}{p_t + p_\epsilon}, p_t + p_\epsilon\right) = V\left(\frac{p_t b_t + p_\epsilon d_t}{p_t + p_\epsilon}, p_t + p_\epsilon\right)$. Therefore, it must be that $\hat{d}_t \in D_t$.

□

Proof of Proposition 3.1. First, consider states where $D_t = R$. Since the mean of the true distribution of d_τ is zero, it is immediate that b_τ will trend toward zero. Alternatively, consider states where D_t is a strict subset of R. Suppose $b_t > 0$. Because D_t is symmetric around $d_t^* \equiv -\frac{p_t b_t}{p_\epsilon} < 0$, it follows that $E[d_t | d_t \in D_t] < 0$. Thus, b_τ will on average trend towards zero in such states. An analogous argument holds when $b_t < 0$. Since b_τ trends towards zero in all states where D_t is non-empty, it follows that $b_\tau \to 0$ as $t \to \infty$.

To see that $D_\tau \to R$, first note that $p_\tau \to \infty$ and therefore $\frac{p_\epsilon d_\tau}{p_\tau + p_\epsilon} \to 0$. That is, because the marginal impact of an outlying data point converges to zero, the firm is less likely to forgo current period revenue to avoid severe belief distortion. Thus, the endpoints of D_t increase in magnitude without bound and ultimately converge to $+\infty$ and $-\infty$.

□

Proof of Lemma 3.3. Consider a continuation payoff where a news firm receives a payoff equal to $\bar{u}$ in every subsequent period. Note that this is strictly higher than the continuation payoff in any plausible state in either the monopoly or duopoly cases. Next, consider the (negative) current period price for which a news firm would be indifferent between reporting news (and receiving the previously mentioned continuation payoff) and not reporting news. That is:

$$P_t + \frac{\delta \bar{u}}{1-\delta} = 0 \tag{C.1}$$

Consider a monopolist news firm and denote its consumers' current bias by $\bar{b}$. Then, we can re-write the above expression as:

$$\chi\gamma \left(\bar{b}^2 + \frac{1}{p_t} + \frac{1}{p_\epsilon} \right) = \frac{\bar{u}}{1-\delta} \tag{C.2}$$

$$\iff \bar{b} = \pm \sqrt{\frac{\bar{u}}{\chi\gamma(1-\delta)} - \frac{1}{p_t} - \frac{1}{p_\epsilon}} \tag{C.3}$$

By construction, given p_t, we know that a monopolist firm when faced with $b_t > |\bar{b}|$ or $b_t < -|\bar{b}|$ will never report news and therefore D_t is empty. Since p_t and p_ϵ are positive numbers, we know that, in any state:

$$-\sqrt{\frac{\bar{u}}{\chi\gamma(1-\delta)}} < \bar{b} < \sqrt{\frac{\bar{u}}{\chi\gamma(1-\delta)}} \tag{C.4}$$

And therefore, for any sequence that does not terminate (i.e. for which D_t is always non-empty), it must be the case that:

$$b_t \in \left(-\sqrt{\frac{\bar{u}}{\chi\gamma(1-\delta)}}, \sqrt{\frac{\bar{u}}{\chi\gamma(1-\delta)}} \right) \tag{C.5}$$

The proof for the duopoly case follows a fortiori since P_t in any given state for a duopolist is weakly lower than P_t in the same state for a monopolist.

□

Proof of Lemma 3.4. By contradiction, suppose that firm 1 serves the whole market (the proof is analogous for the case where firm 2 serves the whole market). There are two possible cases:

- **Case A:** Firm 2 chooses not to report news and firm 1 chooses price $P_1 \leq \bar{u} - \chi\gamma(b_2^2 - v_d) - \phi\gamma(z_1 - b_2)^2$
- **Case B:** Firm 2 reports news but firm 1 chooses price $P_1 < P_2 - 2\phi\gamma\Delta z[b_2 - \bar{z}]$

Note that in both cases, firm 2 earns no revenue AND all consumers consume news, which means that firm 2 does not benefit from any strategic manipulation of consumer beliefs. Thus, if there exists a strategy where firm 2 serves some of the market and earns strictly positive revenue, then the case where firm 2 earns zero revenue cannot be equilibrium.

Note that, because $b_1 \neq b_2$, it cannot be that firm 1 positions itself directly on top of both types of consumers (i.e. it cannot be that $z_1 = b_1$ and $z_1 = b_2$). Thus, firm 2 always has the opportunity to position itself slightly closer than firm 1 to one type of consumer and capture that type while charging a slightly higher price. Since this guarantees positive revenue, it cannot be that one firm serves the whole market in equilibrium.

□

APPENDIX D

A Dynamic Theory of Political Slant: Numerical Computation

D.1. Monopolist's Problem

Reduction to a Finite State Space

We consider a finite-state approximation to the monopolist's problem. In particular, at each iteration we approximate f^i by storing an $m \times n$ matrix $\tilde{f}^i$ defined on an $m \times n$ grid $G = \{(b_i, v_j) : 1 \leq i \leq n, 1 \leq j \leq n\}$ of the bounded (b, v) space. However, even for a finite grid of news d_t the point in the first term of the maximand

$$\left(\frac{b_t v_\epsilon + d v_t}{v_\epsilon + v_t}, \frac{v_t v_\epsilon}{v_t + v_\epsilon}\right) \tag{D.1}$$

may fall outside of the grid G, and so we approximate f^i at such points by simplicial interpolation [Algorithm 6.5, Judd (1998)].

Discretizing News

An additional complication arises with the continuity of news, namely that $d_t \sim \mathcal{N}(0, v_\epsilon + v_\omega)$. Not only are we restricted to computing the decision to report news d_t given a function f^i for a finite set of news,[1] but there is also no closed form solution for the expectation of an arbitrary function of a normally distributed variable, and so we are required to numerically

[1] We could conceivably do this if f is well behaved, in which case we can numerically solve for the zeros of the function and summarize the reporting decision by the finite set of thresholds.

integrate $E_{d_t}[\max\{\dots\}]$. We therefore discretize the space of news as follows. Choose a number $k-1$ of sample points and let $\mathcal{D} = \{d_i : F(d_i) = i/k, i \in (1, k-1)\}$ where F is the cdf of the distribution of d_t. Thus our set of news is sampled equidistantly in terms of the cdf F rather than the news d_t. This simplifies the numeric integration.

Numeric Integration

The integral over possible news is approximated by:

$$\int_{-\infty}^{\infty} h(d)\, dF(d) \approx \sum_{i=1}^{k-1} h(d_{i+j})[F(d_{i+1}) - F(d_i)] \tag{D.2}$$

where $j = 1$ if $F(d_{i+1}) \leq 0.5$ and $j = 0$ otherwise.[2] Note from the way we have defined $\mathcal{D}$ in the previous section we have $[F(d_{i+1}) - F(d_i)] = 1/k$.

Summary

The final implementation is presented in pseudocode form.

(1) Define parameters $\phi, \chi, \gamma, \delta, \bar{u}, v_\omega, v_\epsilon$, discretizing parameters m, n, k, and an error threshold ϵ. In particular, we set $m = n = 100$ and $k = 1000$.

(2) Fix a desired upper bound $\bar{v}$ for v_t and compute $\bar{b}$.

(3) Compute $\mathcal{B} = \left\{-\bar{b} + \frac{i\bar{b}}{n}, i = 0, 1, \dots n\right\}$, $\mathcal{V} = \left\{\frac{i\bar{v}}{m}, i = 0, 1, \dots, m\right\}$ and $\mathcal{G} = \{(b, v) : b \in \mathcal{B}, v \in \mathcal{V}\}$.

(4) Given k, compute $\mathcal{D}$.

(5) Initialize f^0. For simplicity we set $f^0(g) = 0$ for all $g \in \mathcal{G}$.

(6) Define MSE $> \epsilon$ and set $t = 0$.

[2] The idea is to have an approximation that is symmetric across the mean and excludes the endpoints to infinity.

(7) While MSE $> \epsilon$:

- For each $g \in \mathcal{G}$:
  - For each $d \in \mathcal{D}$:
    * Given g and d, compute g^r and g^n, that is for both the case where the monopolist reports d and where he does not.
    * Approximate $f^t(g^r)$ using simplicial interpolation. Note that $g^n = g$ and so $f^t(g^n)$ is already stored.
    * Set:

$$f_d^{t+1}(g) = \max\left\{\bar{u} - \chi\gamma\left(b_t^2 + v_t + v_\epsilon\right) + \delta f^t(g^r), \delta f^t(g^n)\right\}$$

    Numerically integrate $f_d^{t+1}(g)$ over d as proposed in (D.2) to obtain $f^{t+1}(g)$.

- Compute new MSE between f^{t+1}.

(8) Set $V = f^{t+1}$ and f^t and increment t.

D.2. Duopolist's Problem

This section proceeds largely as in the monopoly section, except that the squared dimensionality of the state space $x = (b^1, v^1, b^2, v^2)$ now restricts the granularity of the grid G with which we approximate the value function. Motivated by the results from the monopoly value function, instead of sampling points equidistantly we focus observations in the area around $b = 0$ by letting $\mathcal{B} = \pm\bar{b}/(2^j - 1), j = 1, 2, 3, 4$. We compute $\mathcal{V}$ as previously, but with significantly reduced precision with $m = 8$. The grid $\mathcal{G}$ is then defined as:

$$\mathcal{G} = \{(b^1, v^1, b^2, v^2) : b^1, b^2 \in \mathcal{B}, v^1, v^2 \in \mathcal{V}\}$$

APPENDIX E

A Dynamic Theory of Political Slant: Duopoly Stage Game Algorithm

Full characterization of the stage game equilibrium in the case of duopoly is extremely complex in some states. In order to perform computational analysis, we use the following algorithm to determine how duopolists make slanting and pricing decisions. In some cases, this algorithm gives outcomes exactly equivalent to true equilibrium; in some cases it does not. However, the algorithm is designed to capture many of the same qualitative dynamics demonstrated in Section 3.4.1, and we conjecture that many of the qualitative observations from our results will remain true in the case of formal equilibrium.

E.1. Notation

We will refer to three types of pricing strategies in the algorithm below. Firm i chooses its **monopolist price** P_i^m if it sets its price equal to the total surplus otherwise enjoyed by its consumers. Given competitor price P_j, firm i chooses its **undercut price** $P_i^u(P_j)$ if it charges the price that makes consumer type j indifferent between firms i and j. Given competitor price P_j, firm i charges its **defensive price** $P_i^d(P_j)$ if it charges a price that makes firm j indifferent between remaining at P_j and switching to $P_j^u(P_i^d(P_j))$. A special version of this price are the **mutually defensive prices** $\bar{P}_i^d$ and $\bar{P}_j^d$. This is the pair of prices for which both firms are exactly indifferent between undercutting the other and remaining at their current prices.

We will use the firm's choice of slanting strategy and the firm's choice of position interchangeably. When we claim that firm i is positioned at location z_i, it means that it has chosen the slanting strategy $s_i = s^{z_i}(d)$.

Below are derivations that will be useful:

$$P_1^m = \bar{u} - \chi\gamma(b_1^2 + v_{d,1}) - \phi\gamma(z_1 - b_1)^2 \tag{E.1}$$

$$P_2^m = \bar{u} - \chi\gamma(b_2^2 + v_{d,2}) - \phi\gamma(z_2 - b_2)^2 \tag{E.2}$$

$$P_1^u(P_2) = P_2 - 2\gamma\phi\Delta z(b_2 - \bar{z}) \tag{E.3}$$

$$P_2^u(P_1) = P_1 - 2\gamma\phi\Delta z(\bar{z} - b_1) \tag{E.4}$$

$$P_1^d(P_2^m) = \frac{1}{2}P_2^m + 2\gamma\phi\Delta z(\bar{z} - b_1) \tag{E.5}$$

$$P_2^d(P_1^m) = \frac{1}{2}P_1^m + 2\gamma\phi\Delta z(b_2 - \bar{z}) \tag{E.6}$$

$$\bar{P}_1^d = \frac{4}{3}\phi\gamma\Delta z(b_2 - 2b_1 + \bar{z}) \tag{E.7}$$

$$\bar{P}_2^d = \frac{4}{3}\phi\gamma\Delta z(2b_2 - b_1 - \bar{z}) \tag{E.8}$$

E.2. Algorithm: Pricing

Given z_i and z_j, "equilibrium" prices will be determined as follows:

(1) First, check monopolist prices:

  (a) If neither firm profits from undercutting the other, then each firm selects its monopolist price. Formally, this means $\frac{1}{2}P_i^m \geq P_i^u(P_j^m)$ is true for both firms, and $P_i^* = P_i^m$ for both.

(b) If undercutting is profitable for both firms, then each firm selects its mutually defensive price. Formally, this means $\frac{1}{2}P_i^m < P_i^u(P_j^m)$ for both firms, and $P_i^* = min\{P_i^m, \bar{P}_i^d\}$ for both.

(c) If undercutting is profitable for firm i but not firm j, then check the price pair $[P_i^m, P_j^d(P_i^m)]$.

(i) If firm j still cannot profit from undercutting firm i, then firm i selects its monopolist price and firm j chooses the price that defends against firm i's monopolist price. Formally, this means the following inequalities all hold:

$$\frac{1}{2}P_i^m < P_i^u(P_j^m) \tag{E.9}$$

$$\frac{1}{2}P_j^m \geq P_j^u(P_i^m) \tag{E.10}$$

$$\frac{1}{2}P_j^d(P_i^m) \geq P_j^u(P_i^m) \tag{E.11}$$

Then, $P_i^* = P_i^m$ and $P_j^* = P_j^d(P_i^m)$

(ii) If firm j can now profit from undercutting firm i, then each firm selects its mutually defensive price. Formally, this means the following inequalities all hold:

$$\frac{1}{2}P_i^m < P_i^u(P_j^m) \tag{E.12}$$

$$\frac{1}{2}P_j^m \geq P_j^u(P_i^m) \tag{E.13}$$

$$\frac{1}{2}P_j^d(P_i^m) < P_j^u(P_i^m) \tag{E.14}$$

Then, $P_i^* = min\{P_i^m, \bar{P}_i^*\}$ and $P_j^* = min\{P_j^m, \bar{P}_j^*\}$

E.3. Algorithm: Positioning

Given the process described above that determines prices, "equilibrium" positions are determined as follows:

(1) First, check $z_i = b_i$ for both firms.

(a) If the resulting prices are $[P_i^m, P_j^m]$ from scenario 1(a) in the pricing algorithm or $[P_i^m, P_j^d(P_i^m)]$ from scenario 1(c)(i) in the pricing algorithm, then $z_i^* = b_i^*$ for both firms, and we're done.

(b) If the resulting prices are $[\bar{P}_i^d, \bar{P}_j^d]$ from scenarios 1(b) or 1(c)(ii), then firms position as follows:

$$z_1^* = 2b_1 - b_2 \tag{E.15}$$

$$z_2^* = 2b_2 - b_1 \tag{E.16}$$

And $P_i^* = \bar{P}_i^d$ for both firms, and we're done.

From here, we can substitute the appropriate positions into our prices to get prices as a function of parameters alone, which is sufficient to characterize the full stage game "equilibrium."

APPENDIX F

A Dynamic Theory of Political Slant: Figures

This appendix includes all of the figures referenced in Chapter 3.

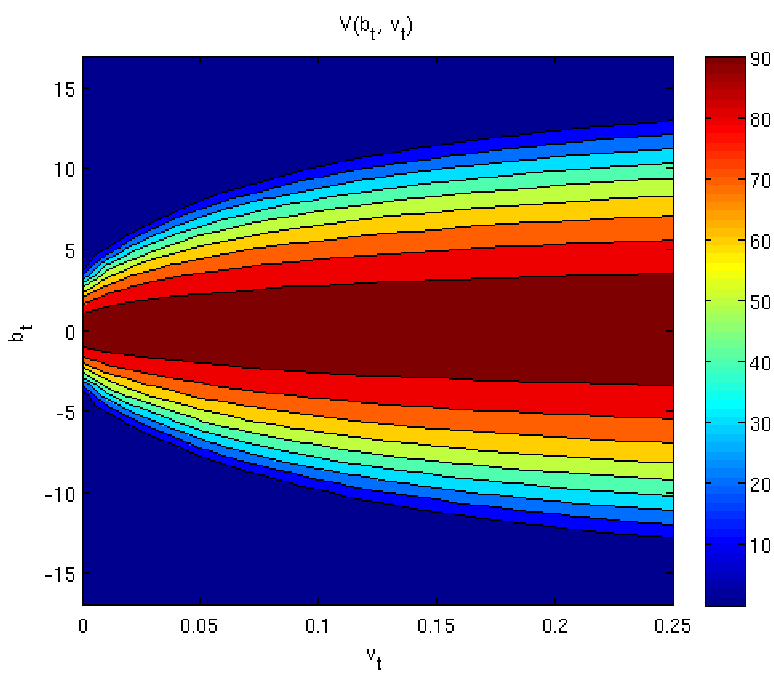

Figure F.1. Numerically computed value function for a monopolist.

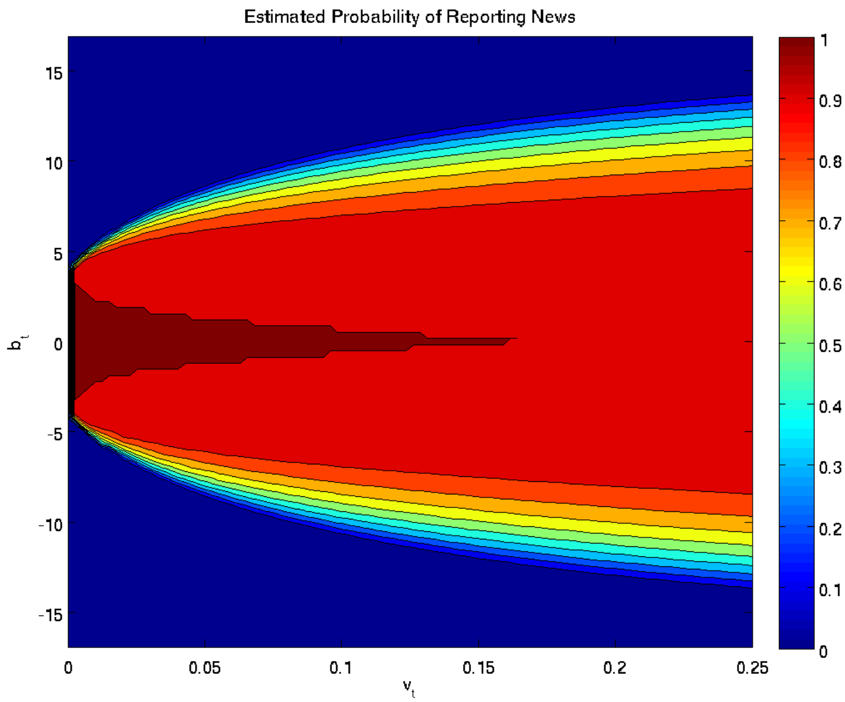

Figure F.2. Probability of revealing news as a function of the state for a monopolist.

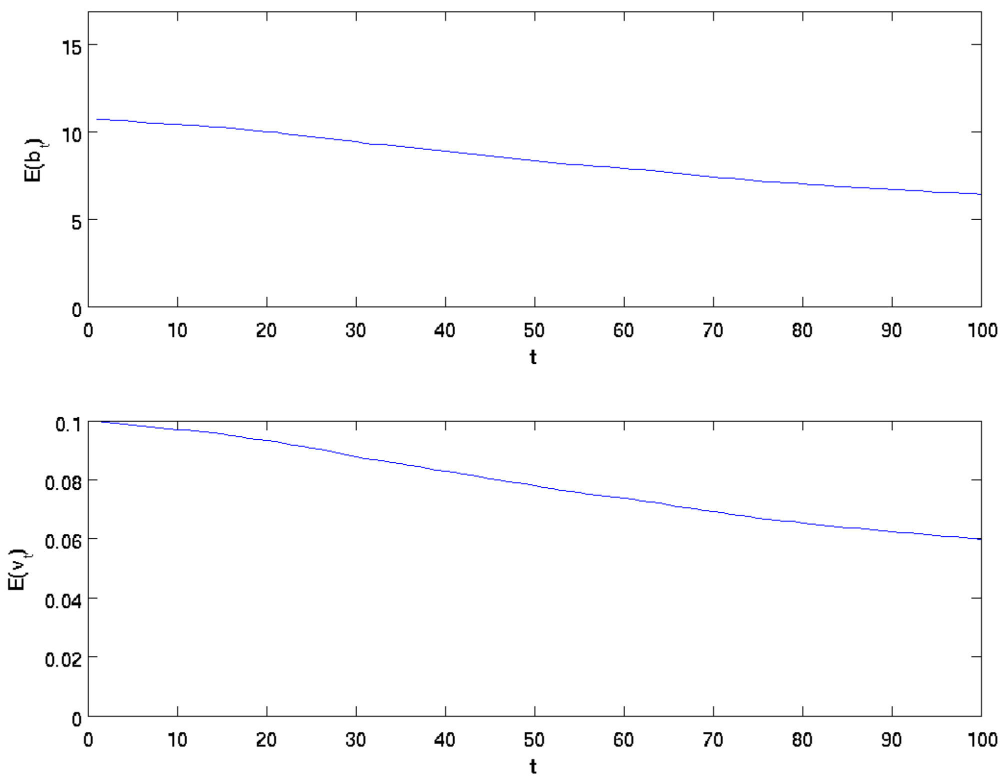

Figure F.3. Simulated path of convergence of (b_t, v_t) with initial condition $(b_0, v_0) = (10.75, 0.10)$ for a monopolist.

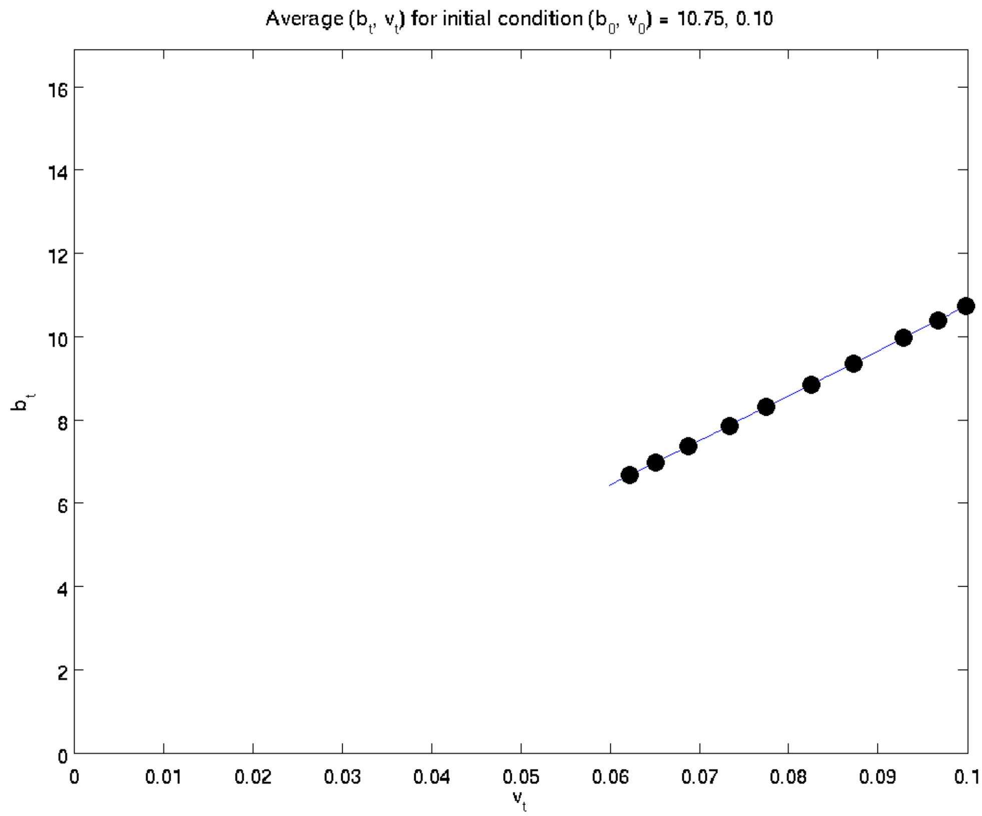

Figure F.4. Simulated rate of convergence of (b_t, v_t) with initial condition $(b_0, v_0) = (10.75, 0.10)$ for a monopolist.

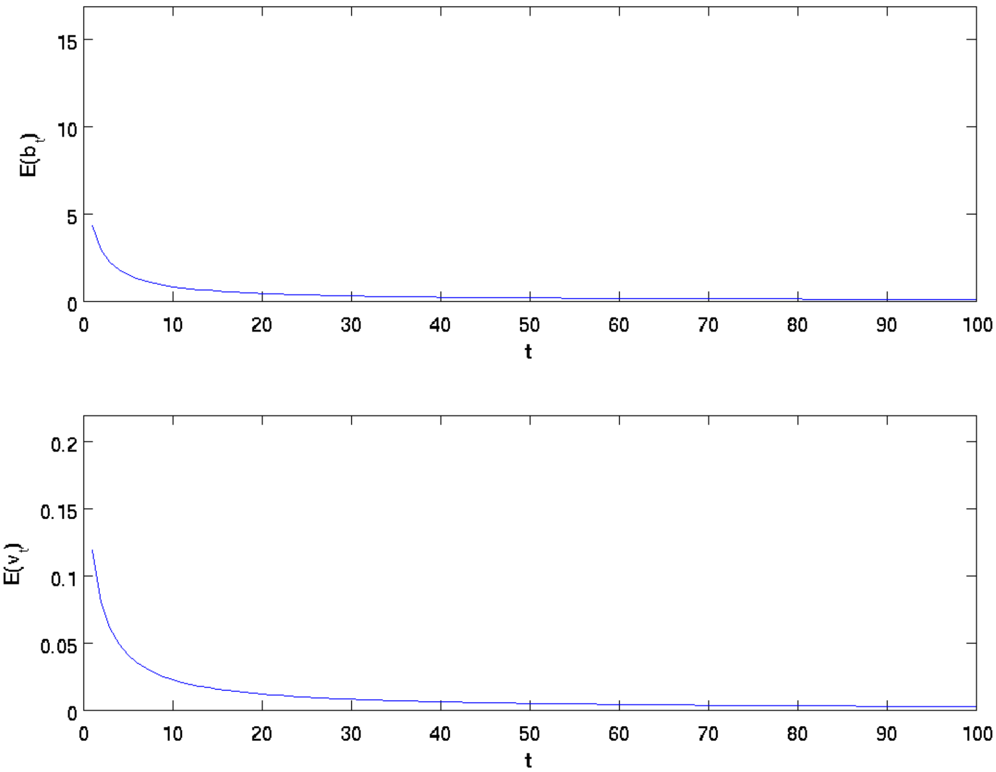

Figure F.5. Simulated path of convergence of (b_t, v_t) with initial condition $(b_0, v_0) = (8.00, 0.22)$ for a monopolist.

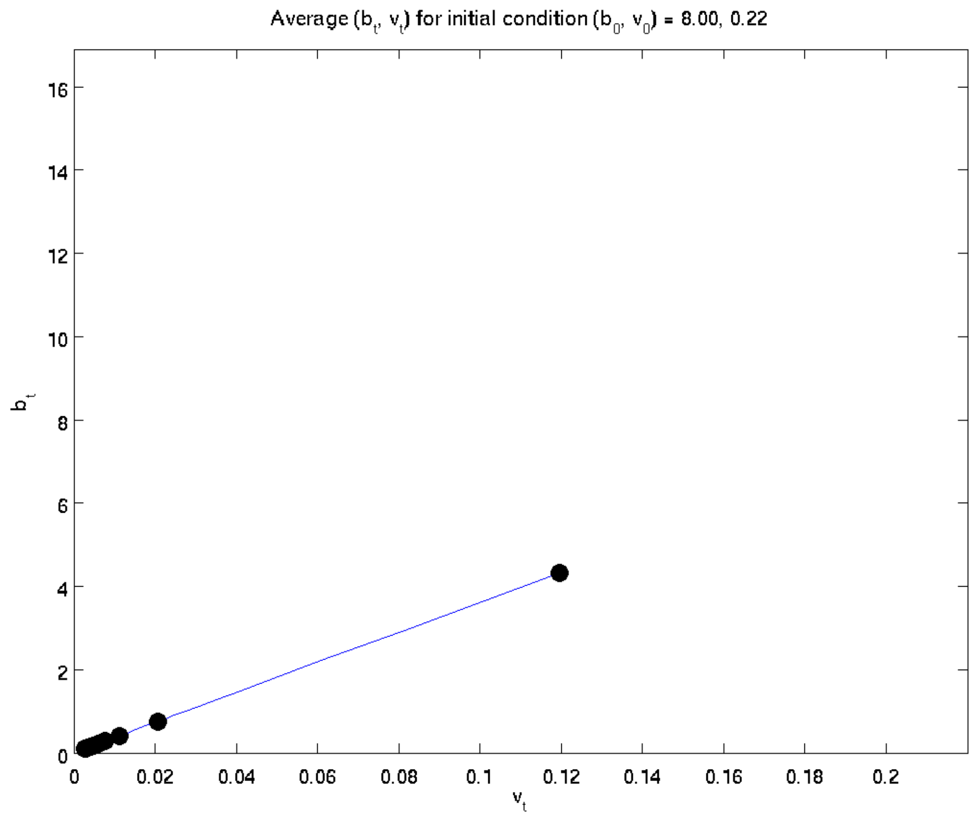

Figure F.6. Simulated rate of convergence of (b_t, v_t) with initial condition $(b_0, v_0) = (8.00, 0.22)$ for a monopolist.

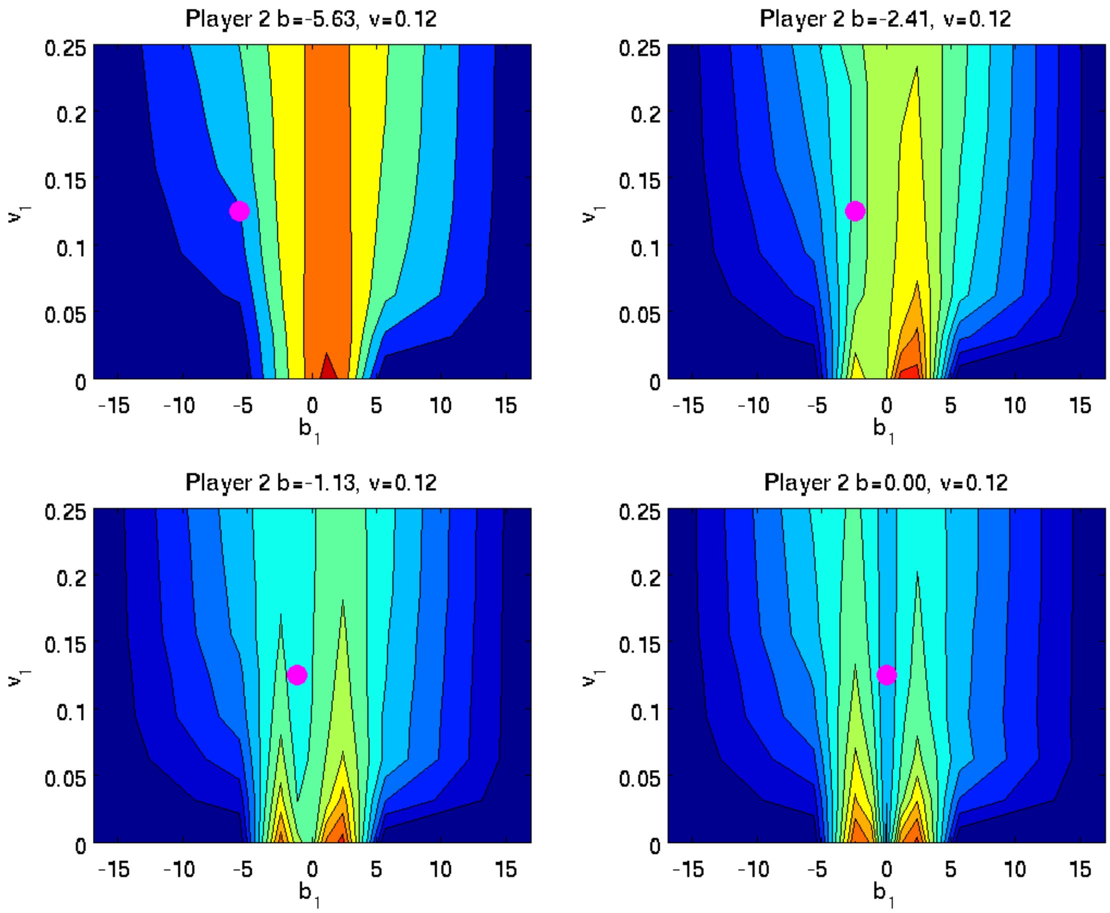

Figure F.7. Value for Firm 1 as we vary b_2.

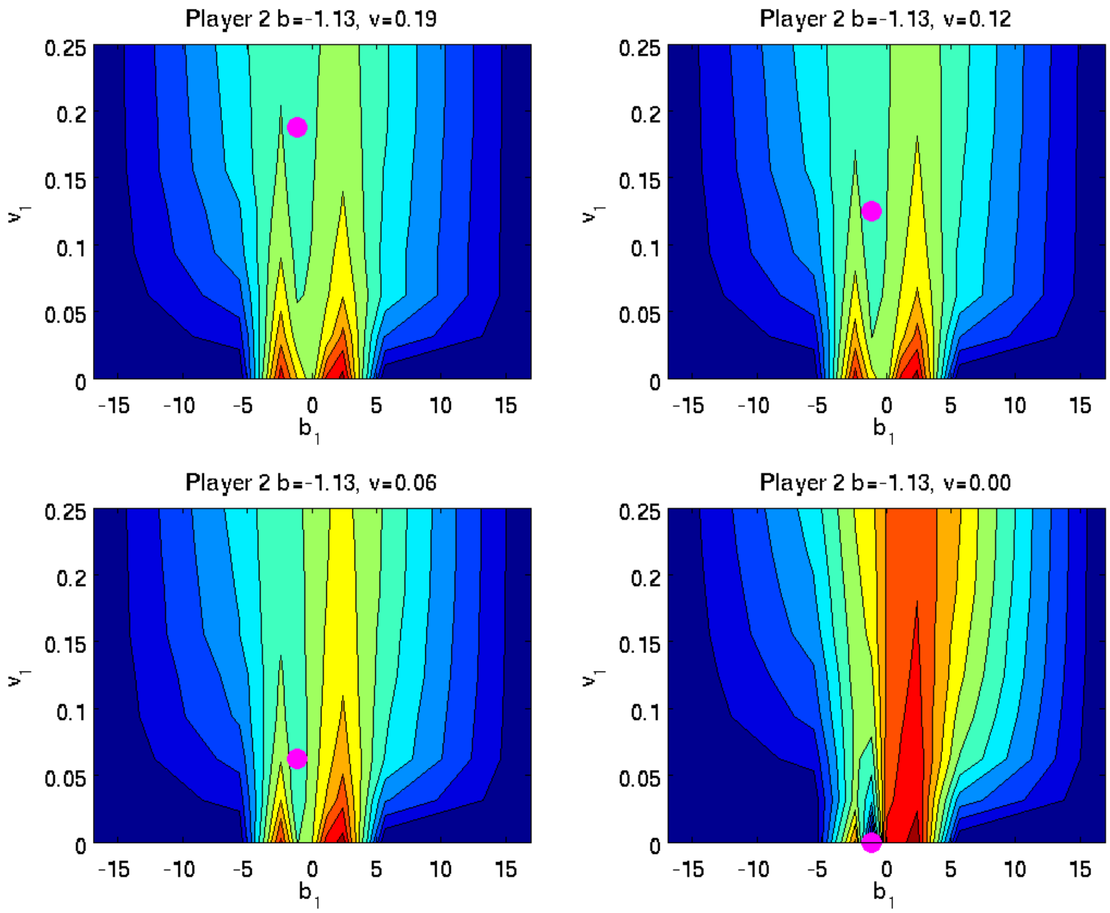

Figure F.8. Value for Firm 1 as we vary v_2.

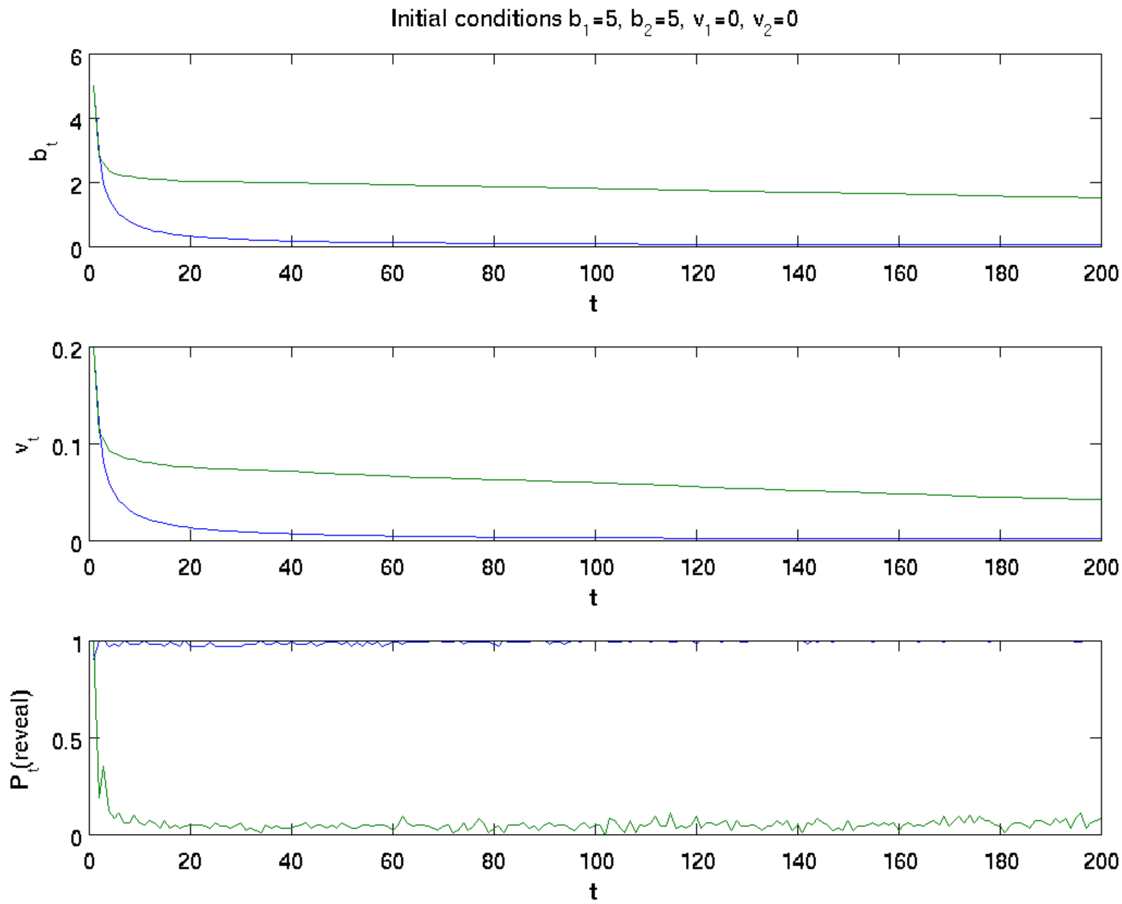

Figure F.9. Average (b_t, v_t) and probability of revelation for duopolists with same initial condition.

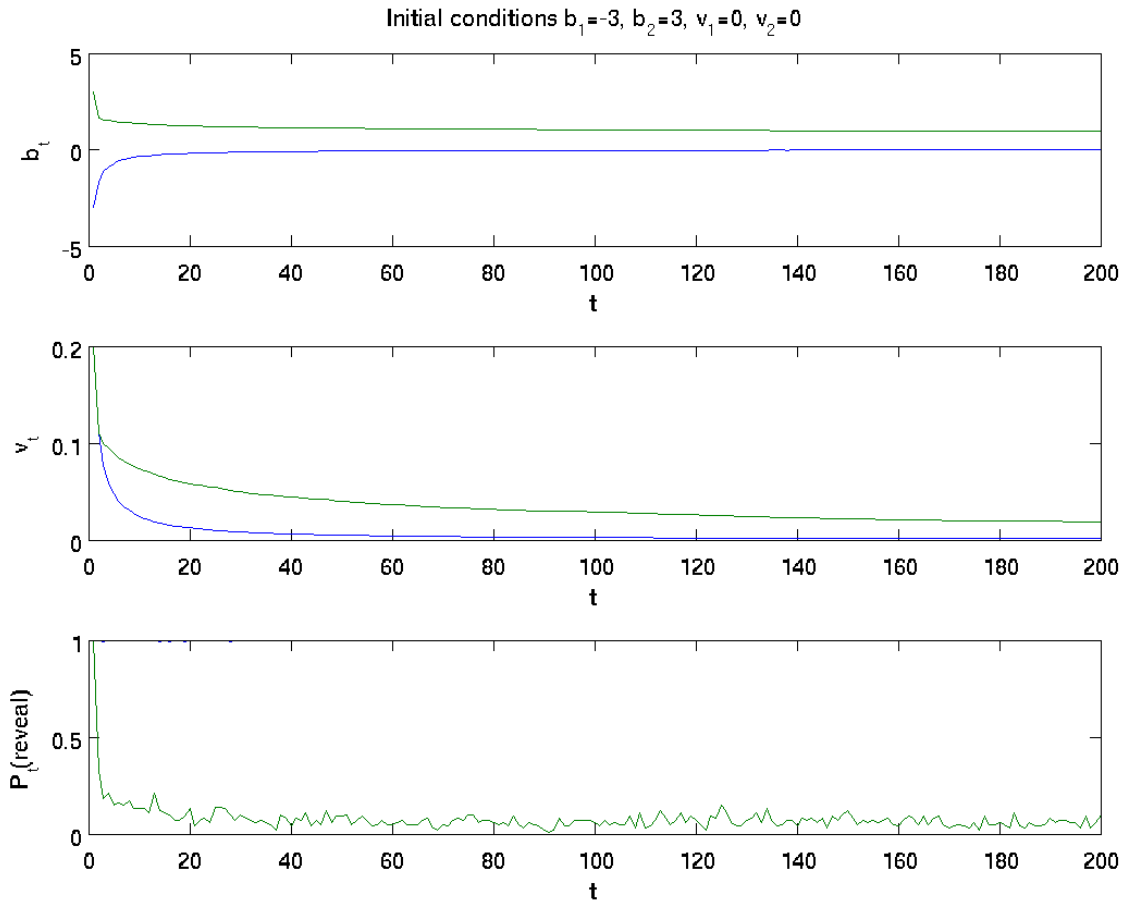

Figure F.10. Average (b_t, v_t) and probability of revelation for duopolists starting at opposite sides of zero.