Revenue from Matching Platforms*

Philip Marx[†] James Schummer[‡]

December 15, 2017

Abstract

We consider revenue maximization for one-to-one matching platforms. Heterogeneous agents from two sides of a market use the platform to form pairs, yielding non-transferable value. The platform commits to a stable matching mechanism and a match-contingent fee for each of the two sides. Despite the fact that agents on the "short" side of the market capture relatively more gross value than those on the long side (when preferences are independently drawn; Ashlagi et al. (2017)), we show that the platform *does not* use relative market sizes to price discriminate across the two sides. The analysis leads to an approximation for the platform's expected revenue through a revenue expression for a *constrained serial dictatorship* mechanism. The approximation shows that the platform's revenue loss from the stability constraint vanishes in large markets. Finally we demonstrate how two types of correlation in preferences lead to two different directions of price discrimination from the baseline case of independence. These effects are absent in classic models of two-sided markets, demonstrating the importance of considering the interaction of capacity constraints and preference correlation.

*This project was partially funded by NSF grant #1534138.

†Education Innovation Laboratory, Harvard University, 1280 Massachusetts Ave, Cambridge, MA 02138. Email: p-marx@kellogg.northwestern.edu.

‡MEDS Department, Kellogg School of Management, Northwestern University. Email: schummer@kellogg.northwestern.edu.

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 the 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 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; 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 ex-

[1] The article "Online Platforms: Nostrums for Rostrums" (*The Economist*, May 28, 2016) observes that for platforms, "established rules of regulation often do not apply."

[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-oen 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.

ample, 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 perceive agents on the other side as homogeneous. 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]

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 stabil-

[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: preferences are heterogeneous on both sides.

[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

ity 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 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.

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 6).

Our first set of results concerning market imbalance appear counterintuitive with respect to the following recent result in the literature on stable matching. In a classic marriage model with independent, uniformly drawn preferences, Ashlagi et al. (2017) show a striking result which can be interpreted as follows. In large markets with 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. As the authors explain,[6] their result is more subtle than the related idea that, in a market of unit-demand buyers and unit-supply sellers of homogeneous objects, a wide range of market (core) prices exist only in the

market, one can interpret these claims as an attempt to determine stable outcomes.

[6] Our explanation here borrows heavily from that of those authors (ibid., p. 72).

knife edge case of a balanced market, e.g. one extra seller depresses prices to cost by creating an "outside option" for each buyer. The extra subtlety is that, in the marriage model, bringing an extra agent to one side of a balanced market creates an outside option that potentially benefits only a *few* agents on the (now) short side. However, this in turn harms a few agents on the long side, in turn benefitting a few more agents on the short side, etc. These effects turn out to ripple through the market, on average benefitting many of the short side's agents.

Following this result one might guess that a monopolist who *controls* the stable matching process could partially capture these unbalanced payoffs by charging unbalanced prices, charging a relatively higher price to the short side of the market. It turns out that this reasoning does not hold for a platform that charges agents in the form of match-contingent fees. 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. (2017). For example, in the particular case that *all* agents' values are drawn i.i.d. from the same distribution, 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, this no-price-discrimination result appears to coincide with results obtained for the many-to-many platform models described in Subsection 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.[7] Therefore it may sound unsurprising when we show, for example, an absence of price discrimination when both sides are equally price sensitive. This is a misleading comparison, however. In many-to-many models, this argument not to price discriminate applies *regardless* of whether agent's values are correlated. Intuitively, in many-to-many models without a capacity constraint, a platform is essentially pricing each transaction separately so correlation plays no role in maximizing expected revenue. 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. Interestingly, we show that the *direction* of price discrimination depends on the type of preference correlation;

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

see Section 6.

Looking at these results in combination, our model sits between the literatures on matching and on two-sided market 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. Yet if market imbalance coexists with *correlation* in agents' values, the platform *does* price discriminate on the basis of market imbalance despite the fact that such discrimination does not occur in many-to-many models without capacity constraints (Rochet and Tirole (2003) and others).

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 depart from the classic Deferred Acceptance algorithm by requiring proposers to "meet" potential mates in some predetermined order. We use a relationship between MAP algorithms and Deferred Acceptance to prove the no-price-discrimination theorem described earlier.

Our analysis of MAP algorithms leads to a closed-form expression that approximates a stable platform's expected revenue when agents have i.i.d. values. This occurs in the special case where the MAP algorithm yields a "constrained-dictatorship" mechanism, representing a platform where agents from one side arrive sequentially and are matched with their favorite remaining (mutually compatible) agent on the other side.[8] By appealing to simulations 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 mere 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.

[8]Though it is not our objective, one could apply our results directly to platforms where agents one one side arrive sequentially to be immediately matched, while all agents on the other side are already present.

1.2 Related literature on two-sided markets

The two-sided markets literature, pioneered by Rochet and Tirole (2003), also considers revenue-maximizing platforms that match two sets of agents who derive value from interacting with each other. In contrast to our approach, many models in this literature exhibit primitives with the following features, making them less relevant for the applications we have in mind.

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

No differentiation: each agent perceives the other side's participants 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 n on the other side of the market. 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 of this leads to one of the fundamental lessons taken from the two-sided markets literature: A profit-maximizing platform should not set prices to the two sides of the markets independently, as if it were pricing two unrelated products. Instead, an increase in the per-transaction price charged to one side of the market should be viewed as a decrease in the marginal cost of providing transactions to the other side of the market, thus affecting that side's optimal price. Furthermore this observation yields the *see-saw effect* (Rochet and Tirole (2006)): if the platform has reason to decrease the price charged to one side of the market (e.g. due to an increase in that side's price sensitivity) then this typically 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

[9]This is the essential payoff structure 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.

market weakly improves stable outcomes for agents on the other side of the market.

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] It is the *all-to-all* feature of these latter models that neutralizes market size effects outright. In our one-to-one model on the other hand, absolute market size *does* impact pricing even when the relative sizes of the two sides does not. Furthermore market imbalance does impact pricing in our model only to the extent that agents' preferences are correlated (see Section 6), an effect not present in standard models of two-sided markets.

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 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 the main points in their setup is that 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 (analogous to one form of correlation we consider in Section 6). Second, their random matching assumption bypasses market size effects in the same way that all-to-all models do; this is natural since their motivation is

[10] An exception is the distantly related work of Ambrus and Argenziano (2009) who consider competing platforms that endogenously create a kind of market size effect. In their equilibrium, one platform ends up being cheaper and larger on one side of the market while the other ends up cheaper and larger on the other side. This equilibrium price discrimination differentiates the platforms. "High value" agents choose their expensive platforms in order to access a greater pool of potential partners on the other side (who have chosen their cheaper platform).

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

the question of market composition or quality. Third, their model contains no same-side externalities: 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)).

Finally, a series of papers considers the impact of various information structures on a platform's mechanism design problem but remains separated from our work by the focus on many-to-many matchings. In this area, Fershtman and Pavan (2016) consider many-to-many matching platforms where private information is persistent over time. Where matchings can change over time, they show that optimal mechanisms are dynamic auctions that determine matchings by applying a scoring rule to reported preferences. Gomes and Pavan (2016) consider the interaction between pricing and matching rules in a many-to-many market where agents have private information about their own values and those of agents on the other side of the market. Jullien and Pavan (2016) study the design of information management policies when agents are uncertain about the participation decisions of other agents; this uncertainty does not apply to our setting of a monopolistic platform that guarantees ex post individual rationality.

2 Model

The agents consist of a finite set of men, $M = \{1, 2, \ldots, M\}$, and a finite set of women, $W = \{1, 2, \ldots, W\}$. We refer to the men's and women's sides of the market as $\mathcal{M}$ and $\mathcal{W}$, respectively. A (one-to-one) **matching** is a function $\mu\colon M \times W \to M \times W$ satisfying the following usual conditions for all $m \in M, w \in W$: (i) $\mu(m) \in W \cup \{m\}$, (ii) $\mu(w) \in M \cup \{w\}$, and (iii) $\mu(m) = w$ if and only if $\mu(w) = m$. We say that agent $i \in M \cup W$ is unmatched (or single) at μ when $\mu(i) = i$.

If man $m \in M$ is matched to woman $w \in W$, m obtains **value** $u_m(w) \in [0, 1]$ and w obtains $u_w(m) \in [0, 1]$. The value of being unmatched is zero (denoted $u_i(i) \equiv 0$ when necessary). These normalizations are not critical to our results. Each value $u_m(w)$ is randomly drawn according to a marginal distribution $F_{\mathcal{M}}$, and each $u_w(m)$ is drawn according to $F_{\mathcal{W}}$, where the corresponding 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 6.[12]

[12]In the independent case, there is zero probability that $u_i(j) = u_i(j')$ for $j \neq j'$. We

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 $m \in M$ makes a payment of $p_{\mathcal{M}}$ to the platform, then his payoff is $u_m(\mu(m)) - p_{\mathcal{M}}$. The symmetric assumption holds for women.

Fixing M and W, we assume that the platform can charge an agent only as a function of (i) to which side of the market that agent belongs (M vs. W), and (ii) whether the agent is matched or single. That is, **prices** are simply a pair of "match-contingent fees" $(p_{\mathcal{M}}, p_{\mathcal{W}}) \in \mathbb{R}^2$, where matched men and women are charged $p_{\mathcal{M}}$ and $p_{\mathcal{W}}$ respectively, while the payments of unmatched agents are normalized to zero. Note that whenever we take M and W as given, the platform is also implicitly setting prices as a function of *market size* (M and W).

2.1 Constraints of the platform

Our main interest is in platforms that charge agents fees to end up as part of a *stable* matching (Gale and Shapley (1962)), i.e. a matching guaranteeing individual rationality and the absence of pairwise blocking. As discussed in the Introduction, stability can be viewed either as a normative criterion or as a positive description of outcomes on decentralized platforms. Regardless of its interpretation and motivation, the classic stability condition becomes endogenous once the platform has the ability to vary prices. This is because an agent who is comparing his current match status with either a departure from the platform (individual rationality) or an alternative partner (pairwise blocking) needs to make this comparison with respect to the platform's prices.

To formalize this, we begin with the individual rationality requirement which states that an agent should not prefer to withdraw from the platform. In our context this means that no matched agent should prefer remaining single *for free* over remaining matched *at the platform's current prices.*[13]

therefore take the standard approach of considering only the realizations of u in which preferences are strict. In Section 6 ties may occur non-trivially, but this turns out to be irrelevant to our results.

[13]This is an *ex post* individual rationality condition, which is stronger than an *ex ante* definition which only requires agents to benefit from the platform in expectation. Most real world platforms would require the stronger definition. This would be true, for instance, on any platform that allows agents to "cancel their reservations," upon learning their match. On the other hand, for platforms content with the weaker (ex ante) condition, the revenue

Definition 1. Fix values u and prices $p = (p_{\mathcal{M}}, p_{\mathcal{W}})$. A matching μ is **individually rational at *p*** when, for all $m \in M$ and $w \in W$,

$$\mu(m) = w \quad \Longrightarrow \quad u_m(w) \geq p_{\mathcal{M}} \text{ and } u_w(m) \geq p_{\mathcal{W}}.$$

The no-blocking-pairs requirement means that no man-woman pair would prefer matching with each other instead of receiving their prescribed matching outcome. In our context we rule out blocking pairs that would prefer matching with each other *while paying the platform's prices in order to do so.*

Definition 2. Fix values u and prices $p = (p_{\mathcal{M}}, p_{\mathcal{W}})$. A matching μ is ***p*-blocked** by man $m \in M$ and woman $w \in W$ when

$$\begin{aligned} &u_m(w) - p_{\mathcal{M}} > u_m(\mu(m)) - p_{\mathcal{M}} * \mathbb{1}_{\mu(m) \in W}, \text{ and} \\ &u_w(m) - p_{\mathcal{W}} > u_w(\mu(w)) - p_{\mathcal{W}} * \mathbb{1}_{\mu(w) \in M} \end{aligned}$$

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

Combining these definitions leads to the following notion of stability.

Definition 3. Fix values u and prices $p = (p_{\mathcal{M}}, p_{\mathcal{W}})$. A matching μ is ***p*-stable** if (i) μ is *individually rational at* p and (ii) μ is not p-blocked by any $m, w \in M \times W$.

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 $p_{\mathcal{M}}$ ($p_{\mathcal{W}}$), 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.[14]

Theorem (Rural Hospital Theorem)**.** *Fix values u and prices $p = (p_{\mathcal{M}}, p_{\mathcal{W}})$. All p-stable matchings contain the same number of marriages.*

---

maximization problem becomes trivial by charging agents their full expected surplus from joining the platform.

[14]Furthermore the set of married agents is constant across all p-stable matchings; see also McVitie and Wilson (1970).

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 observed by the platform, our approach is without loss of generality. Even when preferences are not directly observable by the platform, however, the literature provides various justifications for our approach.

A first justification is that, 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 themselves have complete information about each others' preferences. Specifically, Roth (1984a) shows that in the revelation game induced by Deferred Acceptance, 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 can be made using iterated deletion of dominated strategies (Alcalde (1996)).

Under some of these implementation results, one cannot necessarily predict *which* stable outcome would be obtained. For our purposes, however, the Rural Hospital Theorem makes this irrelevant for the platform concerned simply with revenue-maximization—the same number of marriages are created at any stable matching, so all stable matchings would provide the same revenue to the platform.

A second justification for our approach 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) find that, empirically, the set of stable matchings—and therefore the opportunities for strategic manipulation—are small in the NRMP matching market. Based on a resulting conjecture of theirs, Immorlica and Mahdian (2005) show that, at least whenever preference lists are short, the fraction of participants who can manipulate a stable mechanism vanishes as the market grows large. Consequently, truthful reporting is an approximate equilibrium in such markets (Kojima and Pathak (2009)). Making an additional connection between the ideas of small cores and non-manipulability,[15] Ashlagi et al. (2017) show that in essentially all markets other than perfectly balanced ones, truth telling is an ϵ-equilibrium under any stable mechanism. Technically, none of these results applies directly to our model. As markets grow large, the length of

[15]This idea traces back to Sönmez (1999).

preference lists is exogenously fixed by Immorlica and Mahdian but is of full length in Ashalagi et al. Our model fits "between" these two assumptions: fixing market size, preference lists are endogenously truncated by prices as in Immorlica and Mahdian (2005). However as market size grows (for fixed prices), preference lists grow in proportion to market size (as in Ashlagi et al. (2017)). Therefore as a whole, this body of work justifies our approach of taking realized preferences as given.

A third justification is the empirical evidence (cited in the Introduction) supporting the idea that in certain decentralized markets, real world "equilibrium" outcomes often are stable. For example Hitsch et al. (2010) show that matchings in decentralized, online dating markets are approximately stable even though such platforms clearly cannot dictate which matchings occur. For all three of these reasons, we find our simplifying assumption—that a platform can commit ex ante to producing p-stable outcomes with respect to ex post, realized preferences—to be a plausible simplification of real world platforms.

3 Meet and Propose

A p-stable matching can be found by running the Deferred Acceptance (DA) algorithm (Gale and Shapley (1962)) after the agents' preferences are truncated with respect to prices p. We formalize such an algorithm (DA$_p$) in a somewhat atypical fashion. The reason for this is simply to highlight how the MAP family of algorithms described below relates to DA. To provide the definition, for any prices $p = (p_\mathcal{M}, p_\mathcal{W})$ we say that $(m, w) \in M \times W$ are ***p*-compatible** if $u_m(w) > p_\mathcal{M}$ and $u_w(m) > p_\mathcal{W}$.

Definition 4 (DA$_p$ algorithm)**.** The algorithm takes values u as input and initializes all men to be single. In rounds $t = 1, 2, \ldots$, the following two steps are executed.

Step *t*.1: Each man m who is single "meets" his favorite[16] woman, w, among those to whom he has not already proposed. (If no such women exist, he remains single.) He proposes to w if and only if they are p-compatible.

[16] Ties can be broken arbitrarily. In the case of independent preferences, ties happen with zero probability. With certain kinds of perfect correlation (Section 6) there could be ties, but tie breaking remains irrelevant. The same comments apply to Step t.2.

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. If each man is either matched or has "met" every woman, the algorithm ends; otherwise begin round $t+1$.

It should be clear that DA_p describes the standard DA algorithm, with a bit of redundancy in the notion of men first "meeting" women before proposing. In the standard description of DA in the literature, one dispenses with the notion of "meeting" since men propose only to acceptable women, while women reject offers from unacceptable men. We have intentionally included this redundancy in order to highlight the difference between DA_p and the class of $\mathrm{MAP}_p^{\triangleright}$ algorithms that we use in our analysis.

Specifically, we now introduce algorithms where each man's "meeting order" is not determined by his preferences as it is under the Steps $t.1$ of DA_p, but is determined exogenously. This separates meeting orders from preference orders, though proposals are still tied to preferences when the algorithm determines *whether* a man makes a proposal to a woman he meets. 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.[17]

To describe these algorithms, define a **meeting order** for man $m \in M$, denoted $\triangleright_m$, to be a linear order on W. A profile of meeting orders is denoted $\triangleright = (\triangleright_m)_{m \in M}$.

Definition 5 ($\mathrm{MAP}_p^{\triangleright}$ algorithm)**.** The algorithm is parameterized by a profile of meeting orders $\triangleright$. It takes values u as input and initializes all men to be single. In rounds $t = 1, 2, \ldots$, the following two steps are executed.

Step *t*.1: Each man m who is single "meets" the woman w ranked highest under $\triangleright_m$ among those to whom he has not already proposed. (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. If each man is either matched or has "met" every woman, the algorithm ends; otherwise begin round $t+1$.

[17]Observe that since men make only p-compatible proposals, the women's individual rationality constraints are automatically satisfied.

Each $\text{MAP}_p^{\rhd}$ differs from DA_p to the extent that each $\rhd_m$ differs from m's relative preference order over women according to $u_m()$. If (by chance) the realization of u is such that these orders are equivalent for each man, then the two algorithms produce the same matching. Typically, of course, this does not happen and the algorithms yield different outcomes. Nevertheless, we use $\text{MAP}_p^{\rhd}$ to analyze and approximate the *distribution* of p-stable marriages in a random economy.

As a final observation, note that we have described an "all men propose at the same time" version of both DA_p and $\text{MAP}_p^{\rhd}$. It is well known that DA is invariant to specifications in which men instead propose "one at a time" (e.g. only the lowest-indexed, unmatched man proposes in each round). Similarly it is easily shown that a "one at a time" version of $\text{MAP}_p^{\rhd}$ would produce the same outcome as the algorithm we provide.

4 Independent preferences

We begin with the case of independently drawn values. Formally we suppose that: each $u_m(w)$ is drawn according to $F_{\mathcal{M}}$, each $u_w(m)$ is drawn according to $F_{\mathcal{W}}$, and each of these draws is independent of the others. Throughout Section 4, a **random economy** refers to such independently drawn values (for some fixed M, W).

Our first result establishes a relationship between p-stability and a randomized version of MAP: Running DA_p on a random economy generates the same expected number of marriages as *uniformly randomly* generating a profile of meeting orders, $\rhd$, and then running $\text{MAP}_p^{\rhd}$ on a random economy. In fact, the two processes yield the same *distribution* of marriages, even though their outputs typically differ ex post.[18]

Theorem 1 (random-order MAP marriages $\stackrel{d}{=}$ p-stable marriages)**.** *Fix M, W and prices $p = (p_{\mathcal{M}}, p_{\mathcal{W}})$. Let K_p^{DA} be a random variable representing the number of p-stable marriages in a random economy. Let $K_p^{r\text{MAP}}$ be a random variable representing the number of marriages created under* $\text{MAP}_p^{\rhd}$ *for a random economy when each meeting order $\rhd_m$ is independently drawn from a uniform distribution over all orders. Then K_p^{DA} and $K_p^{r\text{MAP}}$ have the same probability distribution.*

[18]Omitted proofs appear in the Appendix.

The distribution of K_p^{DA} obviously depends on prices $p = (p_{\mathcal{M}}, p_{\mathcal{W}})$. However our next result shows that this is true only to the extent that prices affect the (independent) probability that any given man-woman pair is p-compatible. Recall that (m, w) are p-compatible when $u_m(w) \geq p_{\mathcal{M}}$ and $u_w(m) \geq p_{\mathcal{W}}$. They are thus *incompatible* with probability

$$q(p_{\mathcal{M}}, p_{\mathcal{W}}) \equiv F_{\mathcal{M}}(p_{\mathcal{M}}) + F_{\mathcal{W}}(p_{\mathcal{W}}) - F_{\mathcal{M}}(p_{\mathcal{M}})F_{\mathcal{W}}(p_{\mathcal{W}}) \tag{1}$$

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

Lemma 1 states that, for fixed M, W, and meeting order profile $\triangleright$, the *expected* number of marriages created by $\mathrm{MAP}_p^{\triangleright}$ at prices p is a function only of $q(p)$. That is, $q(p) = q(p')$ implies that $\mathrm{MAP}_p^{\triangleright}$ and $\mathrm{MAP}_{p'}^{\triangleright}$ induce the same expected number of marriages.

Lemma 1 (E[#MAP marriages] is a polynomial function of q)**.** *Fix M, W, and meeting order profile $\triangleright$. For any prices $(p_{\mathcal{M}}, p_{\mathcal{W}})$, let $K_{p_{\mathcal{M}},p_{\mathcal{W}}}^{\triangleright}$ be a random variable representing the number of marriages created under $\mathrm{MAP}_p^{\triangleright}$ for a random economy. Then $E[K_{p_{\mathcal{M}},p_{\mathcal{W}}}^{\triangleright}]$ is a polynomial function of $q(p_{\mathcal{M}}, p_{\mathcal{W}})$. That is, there exists a function $\bar{K}_{\triangleright} \colon \mathbb{R} \to \mathbb{R}$ such that, for all $p_{\mathcal{M}}, p_{\mathcal{W}}$, $E[K_{p_{\mathcal{M}},p_{\mathcal{W}}}^{\triangleright}] = \bar{K}_{\triangleright}(q(p_{\mathcal{M}}, p_{\mathcal{W}}))$, where $q(p_{\mathcal{M}}, p_{\mathcal{W}}) \equiv p_{\mathcal{M}} + p_{\mathcal{W}} - p_{\mathcal{M}}p_{\mathcal{W}}$ is the incompatibility parameter. Furthermore $\bar{K}_{\triangleright}()$ is polynomial in q.*

This lemma follows intuitively from the definition of $\mathrm{MAP}_p^{\triangleright}$. First, each Step $t.1$ of $\mathrm{MAP}_p^{\triangleright}$ looks at prices p only to decide the p-compatibility of each (m, w) pair. *Since meeting orders are exogenous*, each pair's compatibility is randomly determined by a probability independent of the history of the algorithm. Second, each Step $t.2$ of $\mathrm{MAP}_p^{\triangleright}$ determines a woman's favorite man within some set, *conditional on the men in that set having already being deemed compatible*. This determination is (conditionally) independent of prices. Hence the lemma intuitively follows. In fact the proof shows that $\mathrm{MAP}_p^{\triangleright}$ and $\mathrm{MAP}_{p'}^{\triangleright}$ induce the same distribution on number of marriages.

On the other hand this intuition does not apply directly to p-stable marriages and DA_p. In particular, men's meeting orders under DA_p are determined *endogenously* (i.e. by the realization of u), hence it becomes unclear how the decomposition of q into prices $(p_{\mathcal{M}}, p_{\mathcal{W}})$ might affect the distribution of p-stable marriages: the probability of compatibility decreases as a man moves further down his preference list, and hence is not independent of history. As it turns out, however, we show that the conclusion of Lemma 1 does also apply to p-stable marriages, using Theorem 1. If $q(p) = q(p')$, then DA_p and $\mathrm{DA}_{p'}$ induce the same expected number of marriages.

Theorem 2 (E[#p-stable marriages] is a polynomial function of q)**.** *Fix M and W. For any prices $p_{\mathcal{M}}, p_{\mathcal{W}}$, let $K^{\mathrm{DA}}_{p_{\mathcal{M}},p_{\mathcal{W}}}$ be a random variable representing the number of $(p_{\mathcal{M}}, p_{\mathcal{W}})$-stable marriages for a random economy. Then $E[K^{\mathrm{DA}}_{p_{\mathcal{M}},p_{\mathcal{W}}}]$ is a polynomial function of $q(p_{\mathcal{M}}, p_{\mathcal{W}})$. That is, there exists a function $\bar{K}_{\mathrm{DA}}\colon \mathbb{R} \to \mathbb{R}$ such that, for all $p_{\mathcal{M}}, p_{\mathcal{W}}$, $E[K^{\mathrm{DA}}_{p_{\mathcal{M}},p_{\mathcal{W}}}] = \bar{K}_{\mathrm{DA}}(q(p_{\mathcal{M}}, p_{\mathcal{W}}))$, where $q(p_{\mathcal{M}}, p_{\mathcal{W}}) \equiv p_{\mathcal{M}} + p_{\mathcal{W}} - p_{\mathcal{M}}p_{\mathcal{W}}$ is the incompatibility parameter. Furthermore $\bar{K}_{\mathrm{DA}}()$ is polynomial in its argument.*

Proof: Letting $\mathcal{O}$ denote the set of all meeting order profiles, Theorem 1 implies

$$E[K^{\mathrm{DA}}_{p_{\mathcal{M}},p_{\mathcal{W}}}] = E(K^{r\mathrm{MAP}}_{p_{\mathcal{M}},p_{\mathcal{W}}}) = \frac{\sum_{\rhd\in\mathcal{O}} E[K^{\mathrm{MAP}(\rhd)}_{p_{\mathcal{M}},p_{\mathcal{W}}}]}{|\mathcal{O}|}$$

where, by Lemma 1, the numerator is a sum of functions polynomial in $q(p_{\mathcal{M}}, p_{\mathcal{W}})$. □

The takeaway of the theorem is that, under p-stability, a monopolistic platform's demand curve is a function only of the incompatibility parameter $q()$. An immediate consequence of this is that, when the two sides' values are drawn from the same distribution $F = F_{\mathcal{M}} = F_{\mathcal{W}}$, the platform's expected revenue is symmetric in the price vector $(p_{\mathcal{M}}, p_{\mathcal{W}})$. We interpret this as a "non-price-discrimination" result: Despite the fact that matched agents on the short side of a market would extract greater value than those on the long side (as demonstrated by Ashlagi et al. (2017)), the revenue-maximizing platform does not use this fact as a reason to price discriminate against that side of the market. Market imbalance *per se* is not a reason to price discriminate.

Corollary 1 (Revenue symmetry)**.** *Fix M and W, and suppose $F_{\mathcal{M}} = F_{\mathcal{W}}$. For any prices $p = (p_{\mathcal{M}}, p_{\mathcal{W}})$, let $R_{(p_{\mathcal{M}},p_{\mathcal{W}})}$ denote the platform's expected revenue from a random economy under DA_p. Then for all prices p_1, p_2, $R_{(p_1,p_2)} = R_{(p_2,p_1)}$. The same conclusion holds if the platform uses a $\mathrm{MAP}^{\rhd}_p$ mechanism with arbitrary, fixed meeting orders $\rhd$.*

Proof: The planner earns $(p_{\mathcal{M}} + p_{\mathcal{W}})$ from each marriage, so under p-stability, expected revenue has the form $(p_{\mathcal{M}}+p_{\mathcal{W}})\bar{K}_{\mathrm{DA}}(q(p_{\mathcal{M}}, p_{\mathcal{W}}))$ by Theorem 2. Since $q()$ is symmetric in $F_{\mathcal{M}}(p_{\mathcal{M}})$ and $F_{\mathcal{W}}(p_{\mathcal{W}})$ the result follows. The result for $\mathrm{MAP}^{\rhd}_p$ is proven by replacing $\bar{K}_{\mathrm{DA}}$ with $\bar{K}_{\rhd}$ and using Lemma 1. □

We further show that, under a standard hazard rate assumption on $F = F_{\mathcal{M}} = F_{\mathcal{W}}$, revenue-maximizing prices are in fact equal across the two sides, again regardless of the relative size of M in relation to W. To prove this we first observe the intuitive fact that, under either p-stability or $\text{MAP}_p^{\triangleright}$, an increase in the incompatibility parameter q leads to a decrease in the expected number of marriages. For the case of DA, this statement can be proven using a result 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, implying a weakly decreasing number of marriages. Our direct proof in the Appendix also handles the case of $\text{MAP}_p^{\triangleright}$.

Lemma 2 (Marriages decreasing in q)**.** *Fix the* DA_p *mechanism or any arbitrary* $\text{MAP}_p^{\triangleright}$ *mechanism, and denote the expected number of marriages in a random economy by* $\bar{K}(q(p_{\mathcal{M}}, p_{\mathcal{W}}))$*. Then* $\bar{K}()$ *is decreasing in* q*.*

Since the platform's expected revenue is $(p_{\mathcal{M}}+p_{\mathcal{W}})\bar{K}(q(p_{\mathcal{M}}, p_{\mathcal{W}}))$, Lemma 2 demonstrates the typical tradeoff of total price per marriage vs. volume, i.e. $p_T \equiv (p_{\mathcal{M}} + p_{\mathcal{W}})$ vs. $\bar{K}(q(p_{\mathcal{M}}, p_{\mathcal{W}}))$. It is straightforward to observe that the *total* price $p_T = p_{\mathcal{M}}+p_{\mathcal{W}}$ charged by a revenue maximizing platform must be divided amongst the two sides of the market in a way that maximizes $\bar{K}()$, i.e. that minimizes $q()$.[19] Under a standard monotone hazard rate condition, this is typically accomplished with prices that equate hazard rates across the two sides of the market.[20] In the particular case of $F_{\mathcal{M}} = F_{\mathcal{W}}$, this is accomplished by charging the same price to both sides.

Theorem 3 (Monotone hazard rate implies symmetric pricing)**.** *Fix* M *and* W*, and suppose* $F = F_{\mathcal{M}} = F_{\mathcal{W}}$*, where* F*'s hazard rate* $\frac{f(x)}{1-F(x)}$ *is weakly increasing in* x*. Let* $R_{(p_{\mathcal{M}}, p_{\mathcal{W}})}$ *denote the platform's expected revenue from a random economy under a* p*-stable mechanism with prices* $p = (p_{\mathcal{M}}, p_{\mathcal{W}})$*. There exists* $(p^*_{\mathcal{M}}, p^*_{\mathcal{W}})$ *that maximizes* $R_{(p_{\mathcal{M}}, p_{\mathcal{W}})}$ *satisfying* $p^*_{\mathcal{M}} = p^*_{\mathcal{W}}$*.*

The same conclusion holds if the platform uses a $\text{MAP}_p^{\triangleright}$ *mechanism with arbitrary, fixed meeting orders* $\triangleright$*.*

[19]Clearly the choice of how to divide p_T between $p_{\mathcal{M}}$ and $p_{\mathcal{W}}$ affects the platform's revenue through its effect on $q()$. This means that our model fits Rochet and Tirole's (2006) definition of two-sided markets.

[20]The two sides' hazard rates are equated as long as the revenue maximization problem has an interior solution. This may not be the case for arbitrary $F_{\mathcal{M}}, F_{\mathcal{W}}$, and we wish to avoid making further assumptions that add no additional insights. For interested readers, the proof of Theorem 3 provides the FOC that demonstrates this idea.

Obviously when $F_{\mathcal{M}} \neq F_{\mathcal{W}}$, the platform typically charges different prices to the two sides; the intuition to subsidize the price sensitive side demonstrated in models such as Rochet and Tirole (2003) would hold here also. Conversely, if the hazard rate condition does not hold then optimal prices may be unequal even if $F_{\mathcal{M}} = F_{\mathcal{W}}$. To illustrate this as simply as possible, consider the following discrete example.[21]

Example 1 (Optimal, unequal prices)**.** Consider one man and one woman. The value that each agent assigns to its potential mate is (independently) either 0.1 (probability π) or 0.9 (probability $1-\pi$). One can restriction attention to prices $p_{\mathcal{M}}, p_{\mathcal{W}} \in \{0.1, 0.9\}$ and check by inspection that the following price combinations maximize expected revenue.

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

The case $4/9 \leq \pi \leq 4/5$ is the relevant one, demonstrating unequal optimal prices. Note, however, that the *set* of optimal price lists is symmetric in accordance with Corollary 1.

5 Constrained serial dictatorship: $\mathrm{MAP}_p^{=}$

Our next objective is to evaluate a platform's expected revenue under a p-stable mechanism, given arbitrary prices $p = (p_{\mathcal{M}}, p_{\mathcal{W}})$ and for arbitrary distributions $F_{\mathcal{M}}, F_{\mathcal{W}}$. Due to the combinatorial nature of this problem, a tractable expression for this expected revenue remains elusive. On the other hand, it turns out that this revenue would be well-approximated by assuming that the platform instead uses a $\mathrm{MAP}_p^{\triangleright}$ algorithm. In particular we show this approximation to hold in the special case that *all men have identical meeting orders* under $\triangleright$. The advantage of this special case is that it does yield a tractable expression for the expected number of marriages for arbitrary prices and value distributions, and hence can be used to approximate a platform's expected revenue.

[21] The example fails our assumptions but easily demonstrates our point. It can be perturbed into a continuous version that does satisfy our assumptions and yields the same conclusion.

Though we are motivated by this approximation, simulations also suggest that the approximation is in fact a lower bound for the expected revenue of a p-stable platform. This observation is consistent with a related (but logically independent) asymptotic result of Arnosti (2016), who considers matching markets in which one side has "short" preference lists. As markets grow large (while preference lists remain "short") Arnosti shows that the expected number of stable marriages exceeds the expected number of marriages under a random dictatorship mechanism.[22]

For the remainder of this section we consider $\text{MAP}_p^{\rhd}$ algorithms when $\rhd_m = \rhd_{m'}$ for all $m, m' \in M$. Given the ex ante symmetry of the agents in our model, it is without loss of generality to further suppose that each $\rhd_m$ orders the women according to their indices, since if the men had any other (common) meeting ordering, the algorithm would yield the same distribution on the number of marriages. We denote this algorithm by $\text{MAP}_p^{=}$.

$\text{MAP}_p^{=}$ corresponds to a type of *constrained* serial dictatorship mechanism, where each woman sequentially takes a turn choosing her favorite man among the remaining *men who are compatible with her* (if any).[23] It turns out that under $\text{MAP}_p^{=}$, the probability distribution of the number of marriages can be described elegantly. To do this, we introduce the concept of *q-analogs* (i.e. parameterized generalizations) of integers, factorials, and binomial coefficients.

Definition 6. For any real number $q \in [0,1]$, the ***q*-analog of integer *j*** $\in \mathbb{Z}$ is

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

and the ***q*-factorial of *j*** is

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

[22] Arnosti's dictatorship mechanism occurs as a result of what he calls *Common preferences* under Deferred Acceptance. Our conclusions are separate from his, due not only to minor differences in our models, but because we observe this bound across *all* market sizes, not just asymptotically.

[23] Though it is not our objective, one can interpret this algorithm as literally describing platforms where agents on one side (women) arrive randomly over time and are assigned a favorite (but compatible) agent among those remaining on the other side.

The q**-binomial coefficient for integers** $\boldsymbol{k,n} \in \mathbb{Z}_+$ ($k \le n$) 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 distribution of marriages under $\mathrm{MAP}_p^=$ is as follows.[24]

Theorem 4 (Distribution of MAP marriages for identical orders)**.** *Fix M, W, and prices $p_{\mathcal{M}}, p_{\mathcal{W}}$ with incompatibility parameter $q = q(p_{\mathcal{M}}, p_{\mathcal{W}})$. Let $K^=$ be a random variable representing the number of marriages in a random economy created under $\mathrm{MAP}_p^=$. The probability distribution of $K^=$ is given by*

$$P(k; M, W) = (1-q)^k q^{(M-k)(W-k)} \begin{bmatrix} M \\ k \end{bmatrix}_q \begin{bmatrix} W \\ k \end{bmatrix}_q [k]_q! \tag{2}$$

for $0 \le k \le \min\{M, W\}$.

The first two terms in Equation 2 have a straightforward interpretation: fixing k man-woman pairs, $(1-q)^k$ is the probability of mutual compatibility among them, while $q^{(M-k)(W-k)}$ is the probability of mutual incompatibility among all possible pairs of the remaining agents. The remaining q-analog terms in Equation 2 are a probabilistic analog to $\binom{M}{k}\binom{W}{k}k!$, the number of ways to form k man-woman pairs from the market (M, W).

Two special cases of the theorem are 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, an expected-revenue maximizer cares specifically about the expectation of $K^=$ for all intermediate values of q. Fortunately Kemp (1998) provides an expression for this.

Theorem 5 (Kemp (1998))**.** *The expected number of marriages under $\mathrm{MAP}_p^=$ is*

$$E(K^=) = \sum_{j=1}^{\min\{M,W\}} \tfrac{[(1-q^M)\cdots(1-q^{M-j+1})][(1-q^W)\cdots(1-q^{W-j+1})]}{1-q^j} \tag{3}$$

Figure 1 graphs Equation 3 for various levels of q. Fixing the number of men, $M = 50$, we vary the number of women W (x-axis) in order to vary the degree of market imbalance. The figure shows, for instance, that when

[24] Equation 2, known as the absorption distribution, was first described by Blomqvist (1952). Kemp (1998) finds its moments and shows that it is log-concave. Ebrahimy and Shimer (2010) use this distribution to describe the employment rate in a stock-flow matching model with heterogeneous workers and jobs.

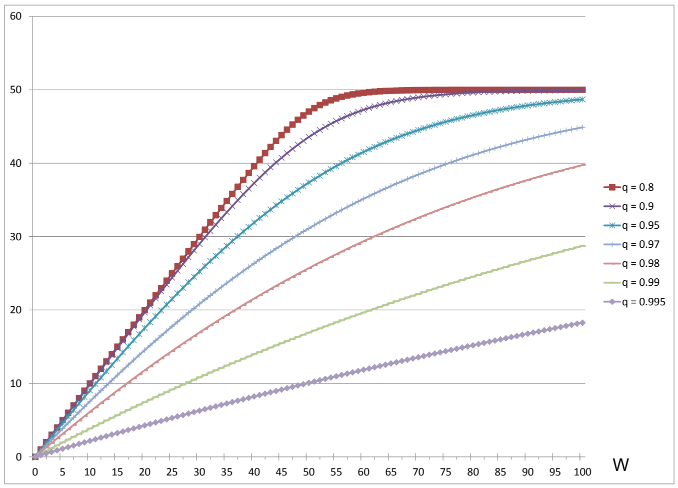

Figure 1. The expected number of marriages under $\mathrm{MAP}_p^=$, varying q and W when $M = 50$.

$q = 0.90$ and $M = W = 50$, there are roughly 43.6 expected marriages under $\mathrm{MAP}_p^=$.[25]

It is clear that, as W grows large, the expected number of marriages converges to 50 since it becomes increasingly likely that each man will be the favorite (remaining) man of some woman in the common meeting order. Consequently, as markets become very imbalanced, the platform can charge relatively higher prices to both sides (high q) yet still expect to create close to the maximum feasible number of possible marriages. Furthermore this convergence happens more quickly for lower incompatibility parameters. That is, when prices are very low, the platform again creates close to the maximum number of possible marriages regardless of the degree of market imbalance. Thus the more interesting cases tend to occur in somewhat balanced markets, where higher prices lead to a non-trivial expected number of single agents.

[25]For example, when values $u_m(\cdot), u_w(\cdot)$ are uniformly distributed on $[0, 1]$, prices $p_{\mathcal{M}} = p_{\mathcal{W}} = 0.684$ yield an incompatibility parameter of roughly $q = 0.90$; this is close to the prices (roughly $p_{\mathcal{M}} = p_{\mathcal{W}} = 0.718$) that would maximize revenue under $\mathrm{MAP}_p^=$.

5.1 $\mathrm{MAP}_p^{=}$ as approximation and bound

We compare the expected number of marriages created under $\mathrm{MAP}_p^{=}$ to that under DA_p, arguing that the former acts as both an approximation and a bound for the latter. As argued above, in sufficiently unbalanced markets (i.e. when the ratio M/W is far from 1) $\mathrm{MAP}_p^{=}$ tends to create close to the maximum feasible number of marriages, namely $\min\{M, W\}$. This observation intuitively extends to DA_p (and for that matter, to all $\mathrm{MAP}_p^{\triangleright}$ algorithms with arbitrary $\triangleright$): even at "high" prices (high incompatibility q), partners can be found for almost all of the agents on the thin side of the market due to the relative thickness of the other side. Therefore it would not be surprising to show that $\mathrm{MAP}_p^{=}$ can be used to approximate the expected number of marriages created by DA_p in the case of unbalanced markets.

Therefore we focus on the "worst case" of balanced markets ($M = W$) where, for a broader range of prices, the expected number of p-stable marriages is not close to $\min\{M, W\}$. We compare (i) the expected number of p-stable marriages (estimated via simulation) to (ii) the expected number of marriages under $\mathrm{MAP}_p^{=}$ (using Equation 3). For a given market size $n = M = W$ and incompatibility parameter q, Figure 2 graphs this difference as a percentage of (i). Our interpretation is that the values in the figure are "small." Since, for the parameters considered in the graph, Equation 3 estimates the expected number of p-stable marriages to within 2.5% of its actual value, Equation 3 serves as a reasonable approximation for the expected number of p-stable marriages even in the worst case of balanced markets. As argued above, this approximation improves in unbalanced markets. An additional observation is that the values in the graph are non-negative, i.e. Equation 3 appears to provide a lower bound for the expected number of p-stable marriages across all market sizes and values of q. As mentioned earlier, this is consistent with a result of Arnosti (2016) in a related model, showing an analogous bound in asymptotically large markets.

The graphs for relatively lower values of q in Figure 2 further suggest that, as the market size grows arbitrarily large, the expected difference between DA_p and $\mathrm{MAP}_p^{=}$ converges to a constant. This is indeed the case. In fact what we show next is that the *absolute* expected number of unmatched agents under $\mathrm{MAP}_p^{=}$ converges to a constant as n grows large.

To give some insight toward the result, consider the probability of achieving a *perfect matching* in a balanced market of size $n = M = W$, i.e. of having zero unmatched agents. The probability that $\mathrm{MAP}_p^{=}$ yields such a matching

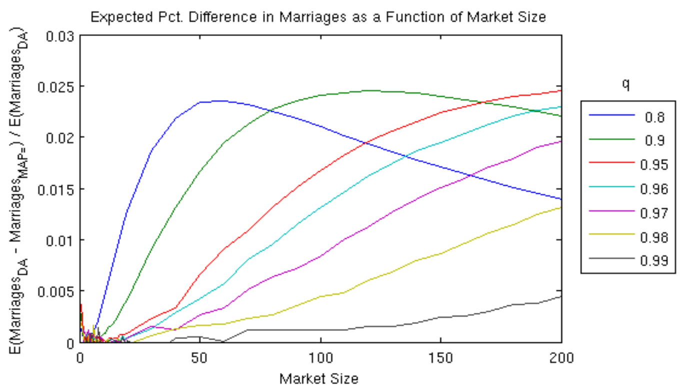

Figure 2. For various balanced market sizes ($M = W \leq 200$) and values of q, the graph shows the percentage by which Equation 3 underestimates the expected number of p-stable marriages; this percentage is at most 2.5% for the parameters represented in the figure. (The unevenness near zero is due to noise from the simulations.)

is given by Equation 2.

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

As n goes to infinity, $P(n;n,n)$ converges to the following expression.[26]

$$\phi(q) \equiv \prod_{i=1}^{\infty}(1-q^i) \tag{4}$$

Therefore $P(n;n,n)$ is bounded away from zero (by $\phi(q)$) across *all* market sizes n.

We can generalize this calculation to find the asymptotic probability of leaving arbitrary, fixed numbers of men and women unmatched, even in unbalanced markets. If there are g single men and h single women, then market

[26]The function ϕ is referred to as *Euler's function*—though unfortunately so are various other functions—and also as a (special case of a) q-Pochhammer symbol. Banerjee and Wilkerson (2016) give a closed-form approximation for ϕ.

sizes must satisfy $M = k + g$ and $W = k + h$ for some number of marriages k. Rewriting Equation 2 in terms of g and h we have

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

Letting the number of marriages $k \to \infty$ (so $M = k + g \to \infty$ and $W = k + h \to \infty$) this converges to the following (see the Proof of Theorem 6).

$$\lim_{k\to\infty} P(k; k+g, k+h) = \phi(q) \frac{q^{gh}}{(1-q)\cdots(1-q^g)\cdot(1-q)\cdots(1-q^h)} \tag{5}$$

Returning to our focus on balanced markets ($g = h$) this implies the following.

Theorem 6 (asymptotic distribution of unmatched agents in balanced markets)**.** *Fix prices $p_{\mathcal{M}}, p_{\mathcal{W}}$ with incompatibility parameter $q \equiv q(p_{\mathcal{M}}, p_{\mathcal{W}}) < 1$, and consider using $\mathrm{MAP}_p^=$ in a random economy. For any positive integer $g \in \mathbb{Z}_{++}$,*

$$\lim_{M=W\to\infty} P(\textit{exactly } g \textit{ men and } g \textit{ women remain single}) \\ = G(g; q) \equiv \phi(q) \frac{q^{g^2}}{((1-q)\cdots(1-q^g))^2} \tag{6}$$

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

Using Equation 6, Figure 3 shows the *expected* number of unmatched agents in arbitrarily large (balanced) markets under $\mathrm{MAP}_p^=$.[27] For example, even when any man-woman pair is incompatible with probability $q = 0.95$, the expected number of single men in arbitrarily large balanced markets is less than 13.4. When $q = 0.80$ the expected value drops below 3.0.

Under $\mathrm{MAP}_p^=$, this observation somewhat trivializes the platform's pricing problem for "very large" markets since, for any fixed prices, unmatched agents become a vanishingly small fraction of the total market size. Therefore a platform can charge close to maximal prices to both sides of the market and yet still create close to the maximum feasible percentage of matches.

[27] Fixing any q, $G(g; q)$ quickly converges to zero as g increases. We approximate $\sum_{g=0}^{\infty} gG(g; q)$ by graphing the sum of the first (sufficiently many) terms.

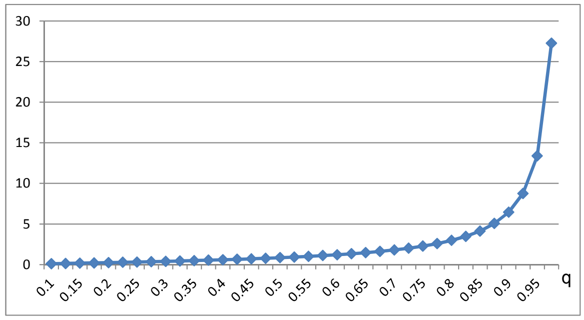

Figure 3. The expected number of unmatched men (equivalently unmatched women) under $\mathrm{MAP}_p^=$ in large, balanced markets as a function of q, for $0.1 \leq q \leq 0.975$.

Nevertheless the pricing problem remains non-trivial for the important case of small- to medium-sized markets.[28] We would argue that many important markets are not large. Even platforms with millions of users, such as Airbnb and Uber, should be viewed as a collection of local markets, each with a relatively small number of participants. At any given time, an Uber driver in North Chicago is not in the same market as a passenger on the city's South Side. An Airbnb renter traveling to New York City on a given night might be interested in staying only in specific areas of that city.

6 Correlated Preferences

We now consider how correlation among agents' values affects the interaction between the platform's prices and the expected number of marriages it creates. We find that correlation alters our price-symmetry results from the independent case (Corollary 1 and Theorem 3), leading a platform to price discriminate between the two sides of the market. The *direction* of price discrimination depends on the *form* of correlation in preferences while the *magnitude* of discrimination depends on the degree of market imbalance. Consequently, the presence of correlation does not validate the false intuition

[28]For a stylistic example, suppose $F_{\mathcal{M}}$ and $F_{\mathcal{W}}$ are uniform distributions on $[0,1]$, and $M = W = 100$. Optimal prices under $\mathrm{MAP}_p^=$ turn out to be approximately $p_{\mathcal{M}} = p_{\mathcal{W}} = 0.77$. Since $q(0.77, 0.77) \approx 0.95$, these prices on average leave roughly 12.6 agents unmatched on each side of the market by Equation 3. Even when $M = W = 500$, optimal prices of roughly 0.86 leave roughly 35 agents unmatched on each side.

(discussed in the Introduction) that a platform should price discriminate by charging relatively higher prices to the short side of the matching market. In fact under one form of preference correlation the platform wants to do precisely the opposite.

Our observation that correlation matters is important for a second reason. It contrasts with the fact that, in the classic two-sided market models[29] where agents can form multiple matches, the platform's pricing decision turns out to be independent of whether agents' preferences are correlated. To elaborate on this, temporarily imagine a many-to-many platform being allowed to price each potential transaction separately, i.e. in our language, imagine that each man-woman pair (m, w) is given *its own* personalized pair of match-contingent prices. With no capacity constraints, the willingness of m and w to match with each other is unaffected by their possibility to be matched with other agents; these two agents will match if and only if they are compatible. Thus the risk-neutral platform maximizes the revenue obtained from this pair by treating (m, w) as an independent pricing problem; optimal *personalized* prices are derived independently of correlation across men's preferences (or across women's preferences). When all men's (resp. women's) values are drawn from the same $F_{\mathcal{M}}$ (resp. $F_{\mathcal{W}}$), the optimal personalized prices charged to (m, w) are the same as those charged to any other pair (m', w'); that is, the constraint to charge *all* men and women the same non-personalized prices does not bind for the risk-neutral platform. This explains why optimal prices in these models are independent of the degree of correlation in values. In contrast that is no longer the case in our model since agents are capacity constrained in forming pairs.

6.1 Two forms of correlation

We separately consider two natural forms of preference correlation in two-sided markets. To describe them, consider the preferences that the women W have over men M. One form of correlation is one where the women tend to agree in their relative preference (i.e. value) for any given man: i.e. the $u_w(m)$'s are correlated across the w's. For example, Airbnb hosts may all tend to think that clean, well-behaved guests are the most desirable ones. Under this form of correlation, agents have a strong common value component in assessing the (heterogeneous) agents on the other side.

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

A second form of correlation is one where any *given* woman's values for various men are correlated. For example, Uber passengers may be indifferent between drivers (at least ex ante, when they commit to being matched). Under this form of correlation, each agent has private values for agents on the other side, but tends to view those agents as being homogeneous. Each agent has his or her own value for participating in the platform relative to some personal outside option, but otherwise is indifferent over partners. It turns out that these two forms of correlation are in a sense dual to each other in a way that is made clear below.

The following definitions define settings where exactly one of the two forms of correlation holds perfectly.[30] In our earlier sections, each $u_i(j)$ was drawn independently (according to either $F_{\mathcal{M}}$ or $F_{\mathcal{W}}$). In this section we assume that the entire profile of values u is drawn from some joint distribution, but under one of the following two sets of assumptions.

Definition 7. Preferences exhibit (perfect) **same-side correlation** when, for any $m, m' \in M$ and $w, w' \in W$, u is drawn in such a way that

- $u_m(w)$ and $u_{m'}(w)$ are perfectly correlated;
- $u_w(m)$ and $u_{w'}(m)$ are perfectly correlated;
- all other pairs of values are independent.

Preferences exhibit (perfect) **cross-side correlation** when, for any $m, m' \in M$ and $w, w' \in W$, u is drawn in such a way that

- $u_m(w)$ and $u_m(w')$ are perfectly correlated;
- $u_w(m)$ and $u_w(m')$ are perfectly correlated;
- all other pairs of values are independent.

An equivalent definition of same-side correlation is as follows. For being matched to a given man m, the women obtain a common value $U_{\mathcal{W}}(m)$ drawn from $F_{\mathcal{W}}$; for being matched to a given woman w, the men obtain a common value $U_{\mathcal{M}}(w)$ drawn from $F_{\mathcal{M}}$. Furthermore these $|M| + |W|$ different values are drawn independently. Similarly, cross-side correlation

[30] Of course most real world applications might involve both of these forms of correlation. Our intention is to analyze how these two forms lead to different pricing behavior, so we consider them separately.

is when each woman w has some "participation value" V_w drawn from $F_{\mathcal{W}}$ obtained from being matched to *any* man, and each man m obtains some V_m drawn from $F_{\mathcal{M}}$ from being matched to *any* woman. Furthermore these $|W| + |M|$ different values are drawn independently.

It turns out that, under either form of correlation, the platform's expected number of marriages under $\text{MAP}_p^=$ is equivalent to that under DA_p.

Theorem 7. *Suppose preferences exhibit either same-side correlation or cross-side correlation. Then for any prices $p = (p_{\mathcal{M}}, p_{\mathcal{W}})$, the expected number of marriages under $\text{MAP}_p^=$ is equal to that under DA_p.*

The simple idea behind the proof of this result illustrates a kind of duality between the two kinds of correlation. Under same-side correlation, the women commonly find any given man acceptable with probability $1 - F_{\mathcal{W}}(p_{\mathcal{W}})$. Thus the total number of "acceptable" men $k_{\mathcal{M}}$ is a binomial random variable, $k_{\mathcal{M}} \sim \mathcal{B}(1 - F_{\mathcal{W}}(p_{\mathcal{W}}), M)$. Similarly the number of acceptable women is a binomial random variable $k_{\mathcal{W}} \sim \mathcal{B}(1 - F_{\mathcal{M}}(p_{\mathcal{M}}), W)$. Under cross-side correlation, a given woman finds *all* of the men acceptable with probability $1 - F_{\mathcal{W}}(p_{\mathcal{W}})$. The number of women "willing to participate" is thus a binomial random variable, $k'_{\mathcal{W}} \sim \mathcal{B}(1 - F_{\mathcal{W}}(p_{\mathcal{W}}), W)$. Similarly the number of willing men is a binomial random variable $k'_{\mathcal{M}} \sim \mathcal{B}(1 - F_{\mathcal{M}}(p_{\mathcal{M}}), M)$.

In the case of same-side correlation, a p-stable matching is obtained by matching $k = \min\{k_{\mathcal{M}}, k_{\mathcal{W}}\}$ agents "assortatively," i.e. with respect to their perceived values. Under $\text{MAP}_p^=$, the $k_{\mathcal{M}}$ acceptable men sequentially propose to (only) the $k_{\mathcal{W}}$ acceptable women (in some order), resulting in $k = \min\{k_{\mathcal{M}}, k_{\mathcal{W}}\}$ marriages regardless of how the women choose amongst their proposals. A similar argument under cross-side correlation yields $k = \min\{k'_{\mathcal{M}}, k'_{\mathcal{W}}\}$ marriages under both algorithms.[31]

The above argument not only establishes an equivalence (in terms of expected marriages) between DA_p and $\text{MAP}_p^=$ but also helps to set up the revenue maximization exercise. In the case of same-side correlation, the platform's problem is to maximize

$$(p_{\mathcal{M}} + p_{\mathcal{W}}) * E(\min\{k_{\mathcal{M}}, k_{\mathcal{W}}\}) \tag{7}$$

where, as above, the distribution of $k_{\mathcal{M}}$ is a function of $(p_{\mathcal{W}}, M)$, and the distribution of $k_{\mathcal{W}}$ is a function of $(p_{\mathcal{M}}, W)$. A similar expression holds

[31]Perfect cross-side correlation yields ties in preferences, but they are easily seen to be irrelevant in the above arguments.

for cross-side correlation, except that $k'_{\mathcal{M}}$ and $k'_{\mathcal{W}}$ depend on $(p_{\mathcal{M}}, M)$ and $(p_{\mathcal{W}}, W)$, respectively. That is, *the form of correlation determines which market size (M vs. W) interacts with which price ($p_{\mathcal{M}}$ vs. $p_{\mathcal{W}}$).* This interaction determines how the platform price discriminates when the market is unbalanced.

To see how, observe that an increase in the size of M shifts the distribution of $k_{\mathcal{M}}$ (or $k'_{\mathcal{M}}$) to the right. On the margin, this shift gives the platform additional incentive to increase whichever price interacts with M (ceteris paribus), e.g. $p_{\mathcal{W}}$ in the case of same-side correlation. Since the distribution of $k_{\mathcal{M}}$ has shifted to the right, it becomes less costly to shift $k_{\mathcal{M}}$ back to the left through a given increase in the interacting price $p_{\mathcal{W}}$. This leads to our conclusion that the platform's level of price discrimination is based on two factors.

(i) The form of preference correlation (same- or cross-side) determines whether the platform price discriminates against the long or short side of the market.

(ii) The degree of market imbalance affects the magnitude of price discrimination.

In general, the revenue maximization problem (Equation 7) is intractable.[32] Therefore we formalize this conclusion by considering a continuous version of our model which abstracts away from the matching frictions of our discrete model.

6.2 Correlation in a large market

Consider our original model, but with a continuum of agents on both sides of the market: a mass $\tilde{M}$ of men and a mass $\tilde{W}$ of women. Our definitions extend to this setting in a straightforward way, so we omit their re-formalization for brevity. The continuum model eliminates the uncertainty in the number (mass) of created marriages, thus eliminating the need to consider an expected value in Equation 7.

For instance, with same-side correlation there are deterministic masses of acceptable men, $\kappa_{\mathcal{M}} = (1 - F_{\mathcal{W}}(p_{\mathcal{W}})) * \tilde{M}$, and of acceptable women, $\kappa_{\mathcal{W}} =$

[32] Part of the intractability comes from the fact that there is no simple formula for the platform's expected volume, i.e. the expectation of the minimum of two arbitrary binomials.

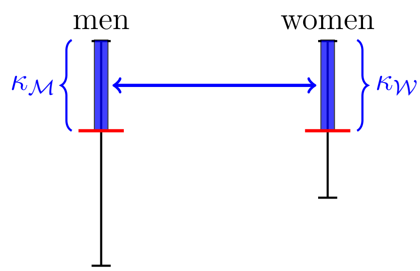

Figure 4. The "best" $\kappa_{\mathcal{M}}$ men and $\kappa_{\mathcal{W}}$ women are matched. Under same-side correlation, $\kappa_{\mathcal{M}}$ is determined by $p_{\mathcal{W}}$ and $\kappa_{\mathcal{W}}$ by $p_{\mathcal{M}}$. If men are the long side of the market as in the figure, then $F_{\mathcal{W}}(p_{\mathcal{W}}) > F_{\mathcal{M}}(p_{\mathcal{M}})$, i.e. a relative premium is charged to the short side of the market. Cross-side correlation flips the relationship between prices and κ's, leading to $F_{\mathcal{W}}(p_{\mathcal{W}}) < F_{\mathcal{M}}(p_{\mathcal{M}})$, i.e. a relative premium is charged to the long side of the market.

$(1 - F_{\mathcal{M}}(p_{\mathcal{M}})) * \tilde{W}$. The mass of p-stable marriages is $\kappa = \min\{\kappa_{\mathcal{M}}, \kappa_{\mathcal{W}}\}$, so the platform typically wants to set prices so that $\kappa_{\mathcal{M}} = \kappa_{\mathcal{W}}$,[33] i.e.

$$\frac{1 - F_{\mathcal{W}}(p_{\mathcal{W}})}{1 - F_{\mathcal{M}}(p_{\mathcal{M}})} = \frac{\tilde{W}}{\tilde{M}} \tag{8}$$

As illustrated in Figure 4 this leads to the conclusion that the short side of the market (say W) is charged a relatively higher price than the long side, M, in the sense that $F_{\mathcal{W}}(p_{\mathcal{W}}) > F_{\mathcal{M}}(p_{\mathcal{M}})$. If $F_{\mathcal{M}} = F_{\mathcal{W}}$ for instance, then the short side of the market is charged a higher price in absolute terms.

In the case of cross-side correlation, the conclusion is reversed. The platform equalizes the two masses of agents "willing to participate" by setting $(1 - F_{\mathcal{W}}(p_{\mathcal{W}})) * \tilde{W} = (1 - F_{\mathcal{M}}(p_{\mathcal{M}})) * \tilde{M}$. This inverts the relationship in Equation 8. Therefore the short long of the market (say M) is charged the relatively higher price: $F_{\mathcal{M}}(p_{\mathcal{M}}) > F_{\mathcal{W}}(p_{\mathcal{W}})$. If $F_{\mathcal{M}} = F_{\mathcal{W}}$, then the long side is charged a higher absolute price.

Finally we relate our conclusions to *changes* in the degree of market imbalance. Consider a ride-sharing platform such as Uber, where preferences approximately exhibit cross-side correlation.[34] This correlation leads to higher

[33] This is true under assumptions that lead to interior solutions to the revenue maximization problem.

[34] Passengers and drivers are roughly indifferent among partners at the time the platform executes a matching, but differ in their willingness to participate. Indeed, an Uber driver's ex ante indifference is enforced by the fact that the driver learns the next destination only *after* agreeing to the next pick up.

relative prices for the long side of the market. Now imagine an increase in the number of passengers. Since the passenger side of the market has become longer, it would be consistent with our results for the platform to adjust its prices in a way that further price discriminates against the long side.[35] Indeed this is what happens under Uber's surge pricing: the passenger-side price is increased while the driver-side price is decreased through higher wages.

On the other hand same-side correlation has the opposite effect. Imagine a platform where agents on any one side tend to agree about the relative desirability of agents on the other side; examples might include dating platforms, or to some extent room rental platforms such as AirBnb. Fixing distributions of values, an increase in the size of one side of the market leads to more value for each member of the other side. E.g., on a (heterosexual) dating site, an increase in the number of agents of one gender, say female, improves the prospects of the males, leading the platform to extract some of this value by increasing the price to the men's side. Of course the situation is more complicated when *both* types of correlation exist simultaneously. Agents on a dating site likely differ in their outside options, corresponding to a form of cross-side correlation. The effect of this would be analogous to Uber's surge pricing: an increase in the number of women leads to lower prices for the men in order to induce their participation.

7 Conclusion

Online platforms have been established—and continue to emerge—in a variety of two-sided markets. While the pricing question for such markets has long been appreciated as important, the literature has not focused on those in which horizontally-differentiated agents from the two sides form exclusive (or capacity-constrained) partnerships. On the other hand, while the broad and growing literature on two-sided matching focuses specifically on such settings, it has not addressed the perspective of a revenue-maximizing platform. We fill this gap by considering a stable matching platform that charges prices to two groups of agents contingent on obtaining a partner on the platform.

Our first main result is a qualitative one that is counterintuitive in light of recent work on stability in unbalanced marriage markets. Ashlagi et al. (2017) show that, with even the slightest imbalance in market sizes, the

[35]Formally showing this would require making additional assumptions on the distributions $F_{\mathcal{M}}$ and $F_{\mathcal{W}}$, going beyond the scope of our discussion.

matched agents on the short side of the marriage market obtain significantly higher payoffs than matched agents on the long side. Thus one might intuitively expect a monopolistic platform to extract some of this surplus imbalance through a corresponding imbalance in prices, charging higher prices to the short side (at least relative to the distributions of values). This turns out not to be the case. For example, when both sides' agents draw i.i.d. valuations for partners from a single distribution, the platform has no incentive to price discriminate against the short side of the market (Corollary 1). For example, if values on both sides of the market are drawn from the same distribution, a standard hazard rate condition induces the platform to price both sides equally (Theorem 3), regardless of the degree of market imbalance.

Our analysis is done by introducing a class of "meet-and-propose" (MAP) algorithms. They differ from Deferred Acceptance by separating the agents' preference orders from the (exogenous) order in which agents make proposals. These algorithms are relevant to our problem because, fixing any prices, the distribution of marriages created by a stable platform (i.e. its "sales volume") turns out to be the same as it would be had the platform instead used a *randomized* MAP algorithm (Theorem 1). Since we can show that MAP algorithms would not induce the platform to price discriminate as above, that first result follows.

Furthermore, the equal-meeting-orders $\mathrm{MAP}_p^=$ algorithm yields a closed form expression for the expected number of marriages, which we demonstrate to be both an approximation and bound for the expected number of marriages for a stable platform. A consequence of this is to trivialize the pricing problem for very large markets with independent preferences: fixing any prices, the number of potential marriages that are prevented by the platform's fees is bounded above by some constant, for all market sizes (Subsection 5.1).

Finally we consider two forms of correlated preferences: same-side (commonality of preferences within the same side) or cross-side (commonality of a given agent's values for agents on the other side). First, when either of these forms of correlation is perfect, the expected number of marriages under the $\mathrm{MAP}_p^=$ algorithm is precisely the same as the expected number of stable marriages. This can be viewed as a robustness check on our earlier use of $\mathrm{MAP}_p^=$ as an approximation for stability. Second, while correlation affects our earlier no-price-discrimination result, we show that the *form* of correlation determines the way in which market imbalance affects the platform's decision to price discriminate.

Intuitively, same-side correlation induces the platform to price discriminate against the short side of the market in order to extract the higher values obtained by those agents. Cross-side correlation induces price discrimination against the long side of the market in order to ration access to the low supply of agents on the other side. Remarkably, in the benchmark case of independent values, these two effects precisely canceled out as mentioned above. These effects highlight the empirical requirements of a revenue-maximizing platform dealing with capacity constrained agents: it needs to estimate not only the distributions of values that individual agents could have, but also the form of correlation these values could have across different agents. Our capacity constrained (one-to-one) model was necessary to establish this observation since, in the classic two-sided market models without capacity constraints, correlation does not affect expected revenue and hence does not affect the platform's incentive to price discriminate.

References

Alcalde, José. "Implementation of stable solutions to marriage problems." Journal of Economic Theory 69.1 (1996): 240-254.

Ambrus, Attila, and Rossella Argenziano. "Asymmetric networks in two-sided markets." American Economic Journal: Microeconomics 1.1 (2009): 17-52.

Armstrong, Mark. "Competition in two-sided markets." The RAND Journal of Economics 37.3 (2006): 668-691.

Arnosti, Nick. "Centralized Clearinghouse Design: A Quantity-Quality Tradeoff" (January 1, 2016). http://dx.doi.org/10.2139/ssrn.2571527

Ashlagi, I., Kanoria, Y., and Leshno Jacob D. "Unbalanced Random Matching Markets: The Stark Effect of Competition." Journal of Politial Economy 125.1 (2017): 69-98.

Banerjee, Shubho, and Blake Wilkerson. "Lambert Series and Q-Functions Near $Q = 1$." (2016). Mimeo. https://arxiv.org/abs/1602.01085.

Blomqvist, N. "On an exhaustion process." Scandinavian Actuarial Journal 35 (1952): 201-210.

Caillaud, Bernard, and Bruno Jullien. "Chicken & Egg: Competition among intermediation service providers." RAND Journal of Economics (2003): 309-328.

Damiano, Ettore, and Hao Li. "Price discrimination and efficient matching." Economic Theory 30.2 (2007): 243-263.

Damiano, Ettore, and Li Hao. "Competing matchmaking." Journal of the European Economic Association 6.4 (2008): 789-818.

Ebrahimy, Ehsan, and Robert Shimer. "Stock-flow matching." Journal of Economic Theory 145(4) (2010): 1325-1353.

Cacioppo, John T., et al. "Marital satisfaction and break-ups differ across on-line and off-line meeting venues." Proceedings of the National Academy of Sciences 110.25 (2013): 10135-10140.

Erdos, Paul, and Alfred Renyi. "On random matrices." Magyar Tud. Akad. Mat. Kutató Int. Közl 8.455-461 (1964): 1964.

Fershtman, Daniel, and Alessandro Pavan. "Dynamic matching auctions." Working Paper (2017).

Gale, David, and Lloyd S. Shapley. "College admissions and the stability of marriage." The American Mathematical Monthly, Vol. 69, No. 1 (Jan. 1962): pp. 9-15.

Gale, David, and Marilda Sotomayor. "Some remarks on the stable matching problem." Discrete Applied Mathematics 11.3 (1985): 223-232.

Gale, David, and Marilda Sotomayor. "Ms. Machiavelli and the Stable Matching Problem." The American Mathematical Monthly 92.4 (1985): 261-268.

Gomes, Renato, and Alessandro Pavan. "Many-to-many matching and price discrimination." Theoretical Economics 11.3 (2016): 1005-1052.

Hitsch, Günter J, Hortaçsu, Ali and Ariely, Dan. "Matching and Sorting in Online Dating." American Economic Review, Vol. 100, No. 1 (2010): pp. 130-163.

Immorlica, Nicole, and Mohammad Mahdian. "Marriage, honesty, and stability." Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics, 2005.

Jullien, Bruno, and Alessandro Pavan. "Information management and pricing in platform markets." Working Paper (2016).

Kara, Tarik, and Tayfun Sönmez. "Nash Implementation of Matching Rules." Journal of Economic Theory 68.2 (1996): 425-439.

Kojima, Fuhito, and Parag A. Pathak. “Incentives and Stability in Large Two-sided Matching Markets.” The American Economic Review 99.3 (2009): 608-627.

Kemp, Adrienne W. “Absorption sampling and the absorption distribution.” Journal of applied probability (1998): 489-494.

McVitie, D. G., and Leslie B. Wilson. “Stable Marriage Assignment for Unequal Sets.” BIT Numerical Mathematics 10.3 (1970): pp. 295-309.

Rochet, Jean-Charles, and Jean Tirole. “Platform competition in two-sided markets.” Journal of the European Economic Association 1.4 (2003): 990-1029.

Rochet, Jean-Charles, and Jean Tirole. “Two-sided markets: a progress report.” The RAND journal of economics 37.3 (2006): 645-667.

Roth, Alvin E. “Misrepresentation and stability in the marriage problem.” Journal of Economic theory 34.2 (1984): 383-387.

Roth, Alvin E.. “The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory.” Journal of Political Ecomony, Vol. 92, No. 6 (Dec., 1984): pp. 991-1016.

Roth, Alvin E. “On the allocation of residents to rural hospitals: a general property of two-sided matching markets.” Econometrica (1986): 425-427.

Roth, Alvin E. “The economist as engineer: Game theory, experimentation, and computation as tools for design economics.” Econometrica 70.4 (2002): 1341-1378.

Roth, Alvin, E., and Elliott Peranson. 1999. “The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design.” American Economic Review, 89(4): 748-780.

Sönmez, Tayfun. “Strategy-Proofness and Essentially Single-Valued Cores.” Econometrica 67 (1999) 677-689.

Weyl, E. Glen. “A price theory of multi-sided platforms.” The American Economic Review 100.4 (2010): 1642-1672.

8 Proofs

Proof of Theorem 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 $p_{\mathcal{M}}$),
(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 $p_{\mathcal{M}}$),
(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[36] hence the two processes are equivalent. □

Proof of Lemma 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 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.[37]
- A temporary match μ which records the set of engagements at the given stage.[38]

[36]We intentionally include this redundancy since it allows derivation of our later results.
[37]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.
[38] Again, this μ can inhabit the same space as a traditional matching function μ.

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:

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

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. 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 $p_{\mathcal{M}} = 0$, $p_{\mathcal{W}} = q$, and $p'_w = q' > q$ (for which $q(p_{\mathcal{M}}, p_{\mathcal{W}}) = q$, $q(p_{\mathcal{M}}, p'_{\mathcal{W}}) = q'$). At $p_{\mathcal{M}} = 0$, 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.[39] 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 $p_{\mathcal{W}}$ and p'_w respectively.

The women's ordinal preferences over men at price $p'_{\mathcal{W}}$ are a truncation of their preferences at $p_{\mathcal{M}}$. It follows from Theorem 2 of Gale and Sotomayor (1985a) that any agent who is unmatched when DA on the women's preferences derived from $p_{\mathcal{W}}$ remains unmatched when running DA on the women's preferences derived from $p'_{\mathcal{W}}$. 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 3. We show that if prices $(p^*_{\mathcal{M}}, p^*_{\mathcal{W}})$ maximize expected revenue, then so does a price of $(p^*_{\mathcal{M}} + p^*_{\mathcal{W}})/2$ charged to both sides. The main steps are written for arbitrary $F_{\mathcal{M}}$, $F_{\mathcal{W}}$ to demonstrate the generality of the idea.

Using Theorem 2 (or Lemma 1 for the case of MAP) and a change of variables, the platform's expected revenue can be written as

$$\max_{p_{\mathcal{M}}, p_{\mathcal{W}}} (p_{\mathcal{M}} + p_{\mathcal{W}})\bar{K}(q(p_{\mathcal{M}}, p_{\mathcal{W}})) = \max_{p_T, p_{\mathcal{M}}} p_T \bar{K}(q(p_{\mathcal{M}}, p_T - p_{\mathcal{M}}))$$

[39]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.

where $p_T \equiv p_{\mathcal{M}} + p_{\mathcal{W}}$ is the total revenue from a single marriage. Let $(p_T^*, p_{\mathcal{M}}^*)$ be a solution to the latter maximization problem. Then

$$p_{\mathcal{M}}^* \in \arg\max_{p_{\mathcal{M}}} p_T^* \bar{K}(q(p_{\mathcal{M}}, p_T^* - p_{\mathcal{M}}))$$

That is, taking p_T^* as given, revenue is maximized by any $p_{\mathcal{M}}$ that maximizes the expected number of marriages $\bar{K}()$. By Lemma 2 this is accomplished by any $p_{\mathcal{M}}$ that minimizes the incompatibility parameter $q(p_{\mathcal{M}}, p_T^* - p_{\mathcal{M}}^*)$. Observe that, since our assumptions guarantee the existence of prices that give positive expected revenue, we must have $q(p_{\mathcal{M}}^*, p_T^* - p_{\mathcal{M}}) < 1$ (otherwise there are zero expected marriages).

Therefore the same must be true for *any* optimal $p_{\mathcal{M}}$, implying both $F_{\mathcal{M}}(p_{\mathcal{M}}) < 1$ and $F_{\mathcal{W}}(p_T^* - p_{\mathcal{M}}) < 1$ (see Equation 1). There is a non-degenerate interval of values for $p_{\mathcal{M}}$ on which these two inequalities are satisfied. Consider the derivative of $q(\cdot)$ with respect to $p_{\mathcal{M}}$, evaluated over this interval.

$$\begin{aligned}
\frac{\partial q(p_{\mathcal{M}}, p_T^* - p_{\mathcal{M}})}{\partial p_{\mathcal{M}}} &= (1 - F_{\mathcal{W}}(p_T^* - p_{\mathcal{M}}))f_{\mathcal{M}}(p_{\mathcal{M}}) - (1 - F_{\mathcal{M}}(p_{\mathcal{M}}))f_{\mathcal{W}}(p_T^* - p_{\mathcal{M}}) \\
&= (1 - F_{\mathcal{M}}(p_{\mathcal{M}}))(1 - F_{\mathcal{W}}(p_T^* - p_{\mathcal{M}})) \left[\frac{f_{\mathcal{M}}(p_{\mathcal{M}})}{1 - F_{\mathcal{M}}(p_{\mathcal{M}})} - \frac{f_{\mathcal{W}}(p_T^* - p_{\mathcal{M}})}{1 - F_{\mathcal{W}}(p_T^* - p_{\mathcal{M}})} \right]
\end{aligned}$$

where the second line avoids division by zero for the candidate values of $p_{\mathcal{M}}$ considered above. Additionally, the sign of $\partial q / \partial p_{\mathcal{M}}$ is determined by the last bracketed term. By the monotone hazard rate condition, this bracketed term is increasing in $p_{\mathcal{M}}$. Hence any $p_{\mathcal{M}}$ for which this term takes the value of zero is optimal. Depending on $F_{\mathcal{M}}$ and $F_{\mathcal{W}}$, such an interior solution may or may not exist. Substituting $F = F_{\mathcal{M}} = F_{\mathcal{W}}$ into the previous equation, however, yields the following signs.

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

Therefore $q(p_{\mathcal{M}}, p_T^* - p_{\mathcal{M}})$ is minimized at $p_{\mathcal{M}} = p_T^*/2$. □

Proof of Theorem 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 Equation 2 holds whenever $M = 1$: the lone man in the economy is either incompatible with each woman ($P(0; 1, W) = q^W$) or not ($P(1; 1, W) = 1 - q^W$).

Using induction on the number of men M, suppose that for any k, Equation 2 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 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 Theorem 6. The theorem results from proving Equation 5. We rewrite Equation 2 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 7. Consider same-side correlation. Fix any prices $p_{\mathcal{M}}, p_{\mathcal{W}}$ and a realization of values u. Given $p_{\mathcal{W}}$, man m is acceptable to each woman if and only if his common value to them exceeds $p_{\mathcal{W}}$; there is some number $k_{\mathcal{M}}$ of such "acceptable" men. Similarly there is some number $k_{\mathcal{W}}$ of acceptable women. Hence any IR matching has at most $k \equiv \min\{k_{\mathcal{M}}, k_{\mathcal{W}}\}$ marriages.

It is simple to see that the unique p-stable matching is the "assortative" one, matching the k "best" men and women, i.e. it contains k marriages.

Consider $\text{MAP}_p^=$. When the men meet the first woman under the common meeting orders $\rhd$ (not necessarily the "best" woman), either she is unacceptable (and none propose), or she is acceptable (and the compatible $k_{\mathcal{M}}$ of them propose). In the latter case, she chooses the best man and rejects the rest. Continuing similarly, each woman either receives no proposals or receives proposals from each of the remaining acceptable men (if any exist). In the latter case, she chooses the best remaining one. Continuing until we run out of either acceptable men or acceptable women, k marriages will result.

A similar argument can be made for cross-side correlation. A difference is this case is that preferences have indifference, but the method of tie-breaking does not impact the expected number of marriages in either algorithm. $\square$