TOO GOOD TO FIRE: NON-ASSORTATIVE MATCHING TO PLAY A DYNAMIC GAME

BENJAMIN SPERISEN AND THOMAS WISEMAN

ABSTRACT. We study stable outcomes in a one-to-one matching market with firms and workers. The model endogenizes how transferable utility is within a match: when a firm-worker pair is matched, they play an infinite-horizon discounted dynamic game with one-sided, observable effort. Partners' types are complementary in determining the maximal feasible payoffs. In our setting, the classic result that with complementary types stable matchings are positively assortative does not hold. Increasing the quality of a match harms dynamic incentives, because a firm cannot credibly threaten to fire a worker who is productive even with low effort. Thus, firms may prefer lower-type workers who will exert more effort. Our analysis suggests a new interpretation of efficiency wages: committing to pay a high wage can improve effort incentives indirectly by making the firm more willing to walk away.

1. Introduction

We consider frictionless matching markets for firms and workers. When a firm-worker pair is matched, they play an infinite-horizon dynamic game. In each period, the worker chooses how much effort to exert. The firm observes the effort level and decides whether to fire the worker or to continue the relationship for the next period. The potential output from a match increases both with the type of the worker (skill or ability) and the type of the firm (technology or capital stock). Further, firm and worker types are complementary – the maximal feasible payoffs are a supermodular function of the partners' types. The actual payoffs that the partners achieve are

Date: December 9, 2016.

We thank V. Bhaskar, Dan Hamermesh, John Hatfield, Ayça Kaya, Leslie Marx, Tomasz Sadzik, Larry Samuelson, Caroline Thomas, and seminar participants for helpful comments.

Sperisen: Department of Economics, Tulane University, bsperise@tulane.edu. Wiseman: Department of Economics, University of Texas at Austin, wiseman@austin.utexas.edu.

determined in equilibrium. Classic results from the literature on two-sided matching (starting with Becker, 1973) show that when payoffs are increasing and supermodular, then stable matchings are positively assortative: high-type workers match with high-type firms. In our setting, that result does not hold. The reason is that increasing the quality of a match harms effort incentives by reducing the firm's willingness to fire. That effect dominates the direct positive effect of complementarity in types, so that some firms prefer lower-type workers who will exert more effort.

The feature that distinguishes our model from the standard matching environment is that here payoffs from a match are determined endogenously. The matching literature typically specifies those payoffs as an exogenous function of types. There are two main branches of that literature. In transferable utility (TU) models, partners can split the payoff generated from their match however they like. In a setting with non-transferable utility (NTU), in contrast, a match generates a specific payoff for each partner. Becker (1973) shows that 1) in the NTU case, stable matchings are positively assortative if each agent's payoff is increasing in his partner's type, and 2) in the TU case, supermodularity of the match value is sufficient for positively assortative matching. Legros and Newman (2007) consider an intermediate case where utility is imperfectly transferable between match partners, and the degree of transferability depends exogenously on types. In our model, transferability is endogenous. Once they are matched, a firm and worker can get any payoffs in the equilibrium set of the dynamic game.

Formally, let the stage-game payoff in a match between a firm (F) of type x and a worker (W) of type y be given by $U^r(e; x, y)$, where $r \in \{F, W\}$ denotes the role of the agent and $e \in [0, 1]$ denotes the level of effort exerted by the worker. Those payoff functions incorporate the market wage rate paid to the worker, which is exogenous but may vary with both x and y. Each payoff $U^r(e; x, y)$ is increasing and supermodular in the types x and y. Effort is costly for the worker but beneficial for the firm, so $U^F(e; x, y)$ is increasing in e and $U^W(e; x, y)$ is decreasing in e. If the firm fires the worker, then both get their type-dependent outside options, $\underline{U}^F(x)$ and $\underline{U}^W(y)$. A key assumption in the model is a single-crossing condition: for low values of x and y, the firm's payoff $U^F(e; x, y)$ exceeds the outside option $\underline{U}^F(x)$ if and only if effort e is high enough, while for high values of x and y, the firm's payoff is strictly greater than the outside option even at zero effort. The consequence is that in matches between high types, the only subgame perfect equilibrium is that the worker exerts zero effort in every period and the firm never fires him: the firm's minimum payoff in the stage game is higher than the payoff from firing. On the other hand, in matches with lower types, the firm prefers to fire the worker if it expects low effort in the future. In that case, firing is a credible threat and effort can be sustained in equilibrium.

That discontinuity in the equilibrium payoff set has implications for stable matching. In evaluating potential partners, a firm faces a potential trade-off between a high-type worker who will not exert effort and a worker with a lower type who will. In particular, the most preferred worker for a firm of type x is either 1) the worker with the highest type y, or 2) the highest-type worker who will exert effort – that is, whose type $y^e(x)$ satisfies $U^F(e; x, y^e(x)) = \underline{U}^F(x)$. As a result, a firm's preferences over workers depend on the firm's type x, even though $U^F(e; x, y)$ is increasing in y for all x and e.

Our findings are as follows. Suppose that the market is thick, in the sense that the distance between an agent's type and the type of the next closest agent is small, and that effort is efficient, in the sense that firms are willing to trade partner type for effort at a higher rate than workers are. Then positively assortative matching is not stable, and a stable, non-assortative matching exists. We also show that the set of stable outcomes has a lattice structure, as in both the TU and NTU cases.

The difficulty that firms face in motivating the most-productive employees is recognized as a problem by human resources professionals. An article on one HR website is titled, "What To Do with the Successful But Lazy Salesperson?:"

> From an organization's perspective, is effort required if performance is there? ... Most companies claim to have goals and systems that require their salespeople to put in the time and effort. In reality, though, I wonder if everyone is just happier to let the sales pro close deals, regardless of effort. (Morris, 2011)

An article on another website describes a related problem:

> Have you ever seen ... an employee who regularly leaves a trail of hate and discontent in his or her wake? This is an employee that regularly upsets coworkers. They are frequently the office bully. Tears are not uncommon sights in people they work with. But ... the boss refuses to even confront them – much less fire them. Why!?!?
>
> [T]he boss considers them "too good to fire". The boss says, "While she may be a Tasmanian devil in the office, that woman could sell ice cubes to an Eskimo. Without her we won't be able to make our third quarter numbers." Or ... "I know he's hard on employees, but he's the only one who knows what he does." (Anonymous, 2011)

If exerting effort consists at least partly of working to cooperate with coworkers, then that scenario also is consistent with our model.

In Section 3, we construct a stable outcome in a parametrized example. Because the payoff functions $U^F(e; x, y)$ and $U^W(e; x, y)$ are supermodular in types, the marginal benefit from matching with a better type of worker increases with the firm's own type: high-type firms care more about their partners' types and less about effort relative to low-type firms. Roughly speaking, then, firms at the top of the type distribution prefer high-type partners, while firms in the middle want to match with lower types to get more effort. The stable matching follows that pattern: positively assortative matching at the top and negatively assortative matching in the middle, where a type-x firm matches with a worker of type $y^e(x)$ (the highest-type worker who will exert effort when matched with the type-x firm). At the bottom, a firm with a very low type x will be able to get effort from any partner, since $U^F(0; x, y)$ will be below the outside option, so it prefers high-type partners, and matching will again be positively assortative.

Thus, the matching between firms and workers is positively assortative at the top and bottom of the type distribution and negatively assortative in the middle. A complication arises, however. Because the function $y^e(x)$ is generically nonlinear, matching firms of type x with workers of type $y^e(x)$ is not feasible – the mass of firms matched would differ from the mass of workers. Instead, it turns out that a stable matching in the middle is very discontinuous. Once firms with types near x "run out" of their most-preferred workers $y^e(x)$ to match with, some of those must match with their "second favorites," whose types are discretely lower.

We also use our model examine to when firms can benefit from paying above-market wages. In particular, suppose that a firm can commit to pay its worker an "efficiency wage": an additional per-period wage as long as the firm does not fire him. If the payoff at the market wage from effort 0 to a type-x firm matched with a type-y worker, $U^F(0; x, y)$, is greater than the firm's outside option $\underline{U}^F(x)$, then the firm cannot credibly threaten to fire. If the firm commits to an additional wage of at least $U^F(0; x, y) - \underline{U}^F(x)$, though, then the firm will be willing to fire, and that threat can be used to enforce effort in equilibrium. Thus, paying an efficiency wage in that situation improves the incentives for effort, but the effect is indirect. The worker would be willing to work to avoid being fired even if the wage were zero. Instead, committing to a positive wage increases the firm's willingness to discipline a worker who deviates. That is, the wage makes effort sustainable not because it makes the threat of being fired more costly for the worker, but because it makes that threat credible for the firm.

Another way for the firm to induce effort when $U^F(0; x, y) > \underline{U}^F(x)$ is by paying a bonus on top of the market wage after any period in which $e = 1$. We find that the firms that benefit the most from paying an efficiency wage, relative to either the market wage or a bonus scheme, are those in industries with the following properties: effort has a high cost for workers but also a high value for the firm, and the value of output is high even when effort is low.

Related literature. As mentioned above, Legros and Newman (2007) examine a matching environment where the extent to which partners can transfer utility to each other is an exogenous function of types. Their condition for stable matchings to be positively assortative, *generalized increasing differences*, fails in our setting. (See Footnote 2 below.) The generalized increasing differences condition requires that the transferability of utility is increasing in types. One interpretation of our contribution is that we provide an economically relevant example where the generalized increasing differences conditions does not hold, so matching is non-assortative.

Two other closely related papers consider matching where the transferability of utility is endogenous. Both Citanna and Chakraborty (2005) and Kaya and Vereshchagina (2015) study settings where partners derive payoffs from a repeated interaction with two-sided moral hazard. Citanna and Chakraborty (2005) find that pairing high-wealth agents with low-wealth partners improves effort incentives. In Kaya and Vereshchagina's (2015) model, effort is unobservable, and an agent's type affects the probability that each period's publicly observed output is high. They find that in some cases negatively assortative matching is stable. In both papers, an agent's type directly affects the ability to provide incentives. Serfes (2005, 2008) and Wright (2004) consider similar issues in models of one-sided moral hazard where a worker's type is his level of risk aversion and the firm's type is the riskiness of its project.

Knuth (1976) and Demange and Gale (1985) prove that the set of stable outcomes has a lattice structure for the cases, respectively, of NTU and TU matching. The argument for the lattice result in our setting combines elements from both proofs.

The literature on partnership games with the possibility of anonymous rematching (including Datta 1996, Kranton 1996a, 1996b, Carmichael and MacLeod 1997, Fujiwara-Greve and Okuno-Fujiwara 2009, and McAdams 2011) is also related. There are two key differences relative to our setting. First, we specify the continuation value from terminating a match as an exogenous function of an agent's type, whereas in partnership games that value is the endogenous expected payoff from rematching. In the conclusion, we briefly discuss the consequences in our model of endogenizing the payoffs after firing along those lines. Second, in most of the literature on partnership games agents are homogeneous – there are no types. Ghosh and Ray (1996) and Watson (1999) are exceptions. In those papers, types are private information, and an agent can learn about his partner's type over time.

The structure of the rest of the paper is as follows: Section 2 presents the model. In Section 3 we analyze a parametrized model in detail. In Section 4 we show how the results from the example generalize. We discuss efficiency wages in Section 5, and in Section 6 we conclude.

2. Model

We consider a one-to-one, two-sided matching market. There are N firms and N workers. Agents are heterogeneous, with types lying in $\left[\underline{\theta}, \bar{\theta}\right] \subseteq \mathbb{R}$. Firm n's type is denoted x_n, where $x_1 < x_2 < \ldots < x_N$. Similarly, worker n's type is y_n, where $y_1 < y_2 < \ldots < y_N$. The sets of agent types are denoted by

$$\mathcal{X}_N \equiv \{x_1, ...x_N\}, \mathcal{Y}_N \equiv \{y_1, ...y_N\}.$$

Some of our results concern the case where the number of agents N grows large. In that case, we focus on sequences $(\mathcal{X}_N, \mathcal{Y}_N)_N$ with the feature that the type distributions for both firms and workers converge weakly to the continuous uniform distribution over $\left[\underline{\theta}, \bar{\theta}\right]$.[1]

Types are publicly observed. When a type-x firm is matched with a type-y worker, they play the following infinite-horizon dynamic game: in each period, the firm decides whether to continue the game or to end it by firing the worker. If it fires the worker, then the continuation payoffs for the firm and for the worker are $\underline{U}^F(x)$ and $\underline{U}^W(y)$, respectively, which are continuous and weakly increasing. If the firm continues the game, then the worker chooses an effort level $e \in [0, 1]$, which is publicly observed. The resulting instantaneous payoffs are $U(e; x, y) = \left(U^F(e; x, y), U^W(e; x, y)\right)$, where $U^F(e; x, y)$ is the firm's payoff and $U^W(e; x, y)$ is the worker's. Payoff functions are continuous and twice continuously differentiable. The payoffs have the following three properties:

First, payoffs are increasing and supermodular in types.

Assumption 1 (*Supermodularity*)**.** *For any* $e \in [0, 1]$*,* $(x, y) \in \left[\underline{\theta}, \bar{\theta}\right]^2$*, and* $r \in \{F, W\}$*,*

$U^r_x(e; x, y) > 0$*,* $U^r_y(e; x, y) > 0$*, and* $U^r_{xy}(e; x, y) > 0$*.*

Second, effort is costly for the worker and beneficial for the firm, the benefit exceeds the cost, and the marginal cost weakly increases and the marginal benefit weakly decreases with the effort level.

[1] Papers on matching with transferable utility that work directly with continuum sets of agents include Gretsky *et al.* (1992), Shimer and Smith (2000), and Board *et al.* (2016).

Assumption 2 (*Efficiency of effort*)**.** *For any $e \in [0,1]$ and $(x,y) \in \left[\underline{\theta}, \bar{\theta}\right]^2$,*
$U^F_e(e;x,y) > 0, U^F_{ee}(e;x,y) \leq 0,$
$U^W_e(e;x,y) < 0, U^W_{ee}(e;x,y) \leq 0,$ *and*
$U^F_e(e;x,y) + U^W_e(e;x,y) > 0.$

Finally, Assumption 3, which relates stage-game payoffs to the continuation payoffs if the worker is fired, is a single-crossing condition. For any effort level and pair of types, workers get a payoff in the stage game strictly higher than the payoff from being fired. Similarly, firms with high enough types do better with any worker type and effort level than the payoff from firing. Low-type firms, on the other hand, get a payoff below their outside option when they are matched with a low-type worker who exerts low effort.

Assumption 3 (*Single crossing*)**.** $U^W(1;x,y) > \underline{U}^W(y)$ *and* $U^F(1;x,y) > \underline{U}^F(x)$ *for all* $(x,y) \in \left[\underline{\theta}, \bar{\theta}\right]^2$*, and*
there exists $\hat{\theta} \in \left(\underline{\theta}, \bar{\theta}\right)$ *such that* $U^F(0;x,x) < \underline{U}^F(x)$ *for all* $x < \hat{\theta}$*, and* $U^F(0;x,x) > \underline{U}^F(x)$ *for all* $x > \hat{\theta}$*.*

Firms and workers discount the future at the common rate $\delta \in (0,1)$ per period, where

$$\delta > \max_{x,y} \frac{U^W(0;x,y) - U^W(1;x,y)}{U^W(0;x,y) - \underline{U}^W(y)} \in (0,1). \tag{2.1}$$

That is, the threat of firing is enough to induce full effort from a worker regardless of match quality. The total payoff from the dynamic game if the worker is fired after T periods and exerts effort e_t in period $t \leq T$ is

$$(1-\delta)\sum_{t=1}^{T} \delta^{t-1} U(e_t;x,y).$$

A public randomization device is available.

Let $E^\delta(x,y) \in \mathbb{R}^2$ denote the set of subgame perfect equilibrium (SPE) payoffs of the dynamic game with a firm of type x and a worker of type y. Given the restriction on the value of the discount factor δ, the equilibrium payoff set has the form described in Lemma 1 below. Define $\underline{e}(x,y)$ as the minimum level of effort that gives the firm a payoff at least as good as its outside option:

$$\underline{e}(x,y) \equiv \text{argmin}_{e\in[0,1]} \left\{ U^F(e;x,y) \geq \underline{U}^F(x) \right\}.$$

Let $\hat{V}(x,y) \equiv \{U(e;x,y) : e \geq \underline{e}(x,y)\}$ denote the set of feasible stage-game payoffs that give the firm a payoff of at least $\underline{U}^F(x)$.

Lemma 1. *The equilibrium set $E^\delta(x,y)$ depends on the values of $U^F(0;x,y)$ and $\underline{U}^F(x)$:*

- *If $U^F(0;x,y) > \underline{U}^F(x)$, then $E^\delta(x,y) = \{U(0;x,y)\}$.*
- *If $U^F(0;x,y) \leq \underline{U}^F(x)$, then $E^\delta(x,y) = co\left\{\hat{V}(x,y), \left(\underline{U}^F(x), \underline{U}^W(y)\right)\right\}$.*

If the firm's payoff when the worker exerts zero effort is strictly higher than the payoff from firing (if $U^F(0;x,y) > \underline{U}^F(x)$), then in equilibrium the firm will never fire the worker. In that case, the only SPE has the worker never exert effort and the firm never fire. If, on the other hand, $U^F(0;x,y) \leq \underline{U}^F(x)$, then there is a SPE where the worker's strategy is to always exert zero effort and the firm's is to always fire. The threat of reverting to that equilibrium can be used to support any effort level in SPE.

This discontinuity in the equilibrium payoff set will generate non-assortative matching. Although the feasible payoffs exhibit complementarity in types, equilibrium payoffs need not.[2]

2.1. **Stability.** In our setting, it is necessary to describe both which firm gets matched with which worker and the play in the dynamic game resulting from each match. An *outcome* consists of a *matching* and an *equilibrium selection rule.* A matching $\mu : \mathcal{X}_N \to \mathcal{Y}_N$ specifies for each type x of firm the type $\mu(x)$ of worker that the firm matches with. An equilibrium selection rule γ maps each matched pair of types $(x, \mu(x))$ to an SPE payoff in $E^\delta(x,\mu(x))$.

We say that an outcome (μ,γ) is *stable* if there is no blocking pair. A firm of type x and a worker of type y form a blocking pair if there is a payoff in $E^\delta(x,y)$ that both firm and worker strictly prefer to the payoffs in their current outcomes. The idea is that a firm (or worker) can ask a worker (or firm) to leave his current partner and

[2]As a result, Legros and Newman's (2007) sufficient condition for stable matchings to be positively assortative, *generalized increasing differences*, fails here.

join it instead, and in making that request can propose an equilibrium to play in the new match. Crucially, the proposal cannot specify an effort level that is not part of an SPE. That is, an outcome includes a fixed SPE for each match, but a potential blocking pair can deviate to any SPE. Note that stability also requires that a matched pair not be able to deviate to a Pareto superior SPE. Thus, the equilibrium selection rule in a stable outcome must specify for each match an SPE with a payoff on the Pareto frontier of $E^{\delta}(x, y)$.

We say that a matching μ is stable if there exists an equilibrium selection rule γ such that the outcome (μ, γ) is stable. We denote the positively assortative matching by μ^{+}: $\mu^{+}(x_n) = y_n$ for all n.

3. Example

In this section, we examine a parametrized example to illustrate our results on stable outcomes. In the example, firm and worker types are distributed across the interval $[0, 2]$. We will focus on the case where N grows large, so that the type distribution approaches a continuous uniform distribution over $[0, 2]$. For expositional purposes, we will sometimes work directly with that limiting distribution.

The continuation payoff after firing is zero for both firms and workers, independent of type: $\underline{U}^{F}(\theta) = \underline{U}^{W}(\theta) = 0$ for all $\theta \in \left[\underline{\theta}, \bar{\theta}\right]$. The stage-game payoff functions are the following:

$$U^{F}(e; x, y) = 2e - 1 + xy$$

and

$$U^{W}(e; x, y) = 2 - e + xy.$$

When the match quality is high enough – specifically, when $xy > 1$ – then the firm will never choose to fire the worker. If $xy \leq 1$, on the other hand, then any effort level below $\underline{e}(x, y) = (1 - xy)/2$ gives the firm a payoff below zero. From 1, the SPE payoff set $E^{\delta}(x, y)$ is given by

$$E^{\delta}(x, y) = \begin{cases} \text{co}\left\{(0,0), (1+xy, 1+xy), (0, \frac{3}{2}(1+xy)\right\} & \text{if } xy \leq 1 \\ (-1+xy, 2+xy) & \text{if } xy > 1. \end{cases}$$

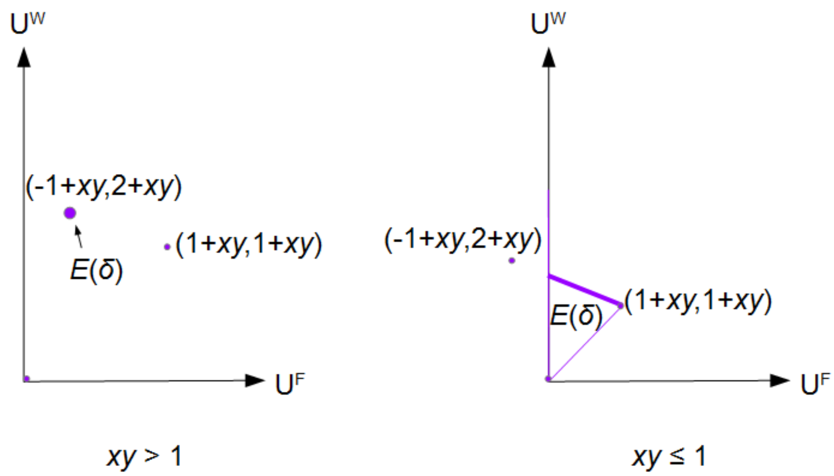

FIGURE 3.1. SPE payoffs as a function of match quality

(Recall that the discount factor δ is high enough – in this case, $\delta > \frac{1}{2}$ – that in any match the threat of being fired is sufficient to induce high effort.) The equilibrium set is illustrated in Figure 3.1.

3.1. **Positively assortative matching is not stable.** In this example, the positively assortative matching μ^+ is not stable when N is large enough – that is, when types are close enough together. Under μ^+, positive effort is sustainable in equilibrium when $x_n y_n \leq 1$. The problem arises at the transition where effort is no longer sustainable. For simplicity, consider the special case where $\mathcal{X}_N = \mathcal{Y}_N$ for all N, so that $x_n = y_n$. Let $x_{n^*(N)}$ be the lowest type of firm strictly higher than 1, and let $y_{n^{**}(N)} < y_{n^*(N)}$ be the highest type of worker such that $x_{n^*(N)} y_{n^{**}(N)} \leq 1$. Then the firm of type-$x_{n^*(N)}$ and the worker of type $y_{n^{**}(N)}$ form a blocking pair as long as N is high enough. Since $x_{n^*(N)} y_{n^*(N)} > 1$, the type-$x_{n^*(N)}$ firm's match partner will not exert effort in equilibrium, and so the firm's payoff will be $-1 + x_{n^*(N)} y_{n^*(N)}$. The firm would prefer to match with the type-$y_{n^{**}(N)}$ worker if that worker exerted effort

e^* strictly above $\underline{e} \equiv \frac{1}{2}x_{n^*(N)}(y_{n^*(N)} - y_{n^{**}(N)})$:

$$2e^* - 1 + x_{n^*(N)}y_{n^{**}(N)} > -1 + x_{n^*(N)}y_{n^*(N)}.$$

The type-$y_{n^{**}(N)}$ worker's current payoff is at most $2+x_{n^{**}(N)}y_{n^{**}(N)}$, so he also prefers to match with the type-$x_{n^*(N)}$ firm and exert effort e^* as long as e^* is strictly below $\bar{e} \equiv y_{n^{**}(N)}(x_{n^*(N)} - x_{n^{**}(N)})$:

$$2 - e^* + x_{n^*(N)}y_{n^{**}(N)} > 2 + x_{n^{**}(N)}y_{n^{**}(N)}.$$

For large N, $\underline{e} < 1$. Further, because $x_{n^*(N)} - x_{n^{**}(N)} = y_{n^*(N)} - y_{n^{**}(N)}$, we have $\bar{e} > \underline{e}$. Since $x_{n^*(N)}y_{n^{**}(N)} \leq 1$, any effort level is sustainable in equilibrium when the type-$x_{n^*(N)}$ firm and the type-$y_{n^{**}(N)}$ worker are matched. Thus, they form a blocking pair.

Intuitively, the reason that the switch benefits both is that firms put a higher utility weight on effort relative to match quality than do workers. The fact that payoffs are complementary in types acts in the opposite direction: the (high) type-$x_{n^*(N)}$ firm cares more about the type of its partner than does the (low) type-$y_{n^{**}(N)}$ worker. If the difference in types is small, though (that is, if N is large), then the first effect dominates: the type-$x_{n^*(N)}$ firm is willing to accept a slightly lower-quality partner in exchange for positive effort, while the type-$y_{n^{**}(N)}$ worker is willing to supply that increased effort for a higher-quality partner.

3.2. **A stable outcome.** Here we focus on the preferences of firms over match partners to derive the properties of stable outcomes for large N.

The firm with the highest type, $x = 2$, prefers to match with the highest type worker, $y = 2$, and get no effort rather than match with the highest type of worker who would exert effort, $y = \frac{1}{2}$. That is, $-1 + 0 + 2 \cdot 2 > -1 + 2 + 2 \cdot \frac{1}{2}$. Since firm $x = 2$ and 0 effort is the most-preferred match result for the worker with type 2, that pairing must be part of any stable outcome. By the same argument, the next highest firm type will match with the next highest worker type, who will exert no effort. That logic extends for every firm with a type above $x = \sqrt{3} \approx 1.73$: because $-1 + x^2 > 2$, those firms prefer zero effort from a worker of their own type to full effort from the

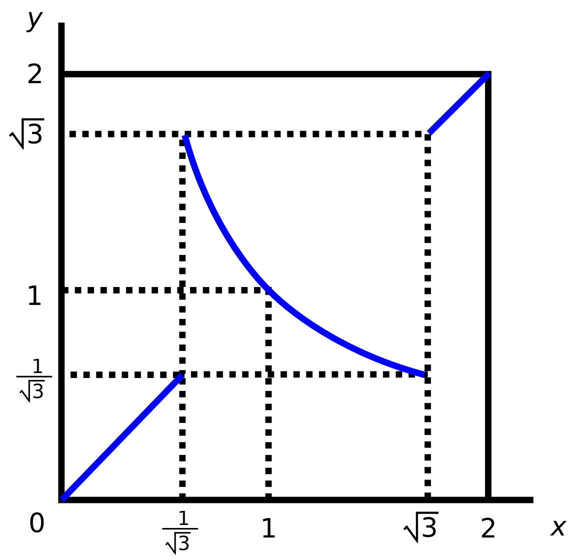

FIGURE 3.2. A non-measure preserving matching

best potential partner willing to work when matched with them, $y = 1/x$. Thus, every stable outcome features positively assortative matching at the top.

Below $x = \sqrt{3}$, firms' preferences depend on equilibrium effort levels. For a start, we look for a *within-match firm-optimal* outcome $(\tilde{\mu}, \gamma^{\max})$, where the equilibrium selection rule $\gamma^{\max}$ specifies the highest effort level consistent with equilibrium for each match in μ. That is, if $xy \leq 1$, then the worker exerts effort 1, yielding payoffs $(1 + xy, 1 + xy)$. If $xy > 1$, then the worker puts in 0 effort, and payoffs are $(-1 + xy, 2 + xy)$.

In that case, a firm with a type just below $x = \sqrt{3}$ prefers being matched with the highest type of worker who will exert effort, $y = 1/x$, rather than getting no effort from the best worker type remaining, $y = \sqrt{3}$. The same argument applies for firm types down to $x = 1/\sqrt{3} \approx 0.58$. In that range there is *negatively* assortative matching as a firm of type x matches with a worker of type $1/x$. Below $x = 1/\sqrt{3}$, a firm's type is low enough that effort will be attainable ($xy < 1$) when the firm is matched with any type of worker remaining, and so matching is positively assortative again at the bottom. Figure 3.2 illustrates that matching.

The problem with the matching just described, however, is that the continuous limit (that is, as N grows without bound) of the matching for middle types – where a type x firm matches with a type $1/x$ worker – is not measure preserving. It specifies, for instance, that firms with types between 1 and $\sqrt{3}$ (measure $\frac{1}{2}(\sqrt{3}-1) \approx 0.37$) are matched to the workers with types between 1 and $1/\sqrt{3}$ (measure $\frac{1}{2}(1-1/\sqrt{3}) \approx 0.21$). Too many firms are matched to too few workers. Similarly, the relatively few firms with types between 1 and $1/\sqrt{3}$ are supposed to match with the relatively many workers with types between 1 and $\sqrt{3}$. For a large but fixed N, the problem is that for a firm of type x, its ideal type $y = 1/x$ of worker may not be an element of $\mathcal{Y}_N$. In that case, the firm's most preferred existing worker will be

$$\max\left\{y \in \mathcal{Y}_N : y \leq \frac{1}{x}\right\}.$$

But then multiple firms will have the same highest worker type who will exert effort for them.[3] Thus, it is not feasible to match each firm with the highest worker type who will exert effort.

The solution is that in the region where the firms outnumber their most preferred workers (that is, firm types in $[1, \sqrt{3}]$), some of the firms must match with their second choices instead. In this "high middle" region, first match the firms with their preferred worker types $1/x$ to the extent possible. There will be a leftover mass of $\frac{1}{2}(\sqrt{3}-1) - \frac{1}{2}(1-1/\sqrt{3}) \approx 0.15$ firms still unmatched. Those firms must match with their second-favorite workers. For example, consider a firm whose type x is just below $\sqrt{3}$. If the firm cannot get its preferred worker (among those workers remaining after firms with higher types have matched), type $1/x$ just above $1/\sqrt{3}$, then its next choice will be either a slightly lower type or a much higher type. (Recall that worker types between $y = 1$ and $y = \sqrt{3}$ are still available, as are types below $1/\sqrt{3}$.) The firm would never choose a type y *slightly* higher than $1/x$ – that worker would not exert effort, and conditional on no effort the firm would rather match with an even higher type. Suppose that the firm prefers the slightly lower type. Then it will match with a worker of type just below $1/\sqrt{3}$; call that type $\hat{y}(x)$.

Next, a firm with a slightly lower type x' faces a similar choice. Suppose that it also, rather than matching with a much higher type, prefers to match with the highest remaining worker who will exert effort, which now is the type just below

[3] For example, suppose that types lie in $\{0.1, 0.2, \ldots, 1.9, 2\}$. Then for a firm of type 1.3, the largest y such that $1.3y \leq 1$ is $y = 0.7$. But $y = 0.7$ is also the largest y such that $1.4y \leq 1$.

$\hat{y}(x)$, $\hat{y}(x')$. Suppose that all the leftover firms with types from $x = \sqrt{3}$ down to $x = \sqrt{3} - \Delta$ for some small $\Delta > 0$ match in that fashion. Then overall, the measure $\frac{1}{2}\Delta$ of firms with types in that interval (including those who matched in the "first round" with worker type $1/x$) match with the workers of types below $1/(\sqrt{3} - \Delta)$ and above $\hat{y}(\sqrt{3} - \Delta)$, because $1/x$ is decreasing in x and $\hat{y}(x)$, by construction, is increasing in x. For the measure of workers to match the measure of firms, it must be that $1/(\sqrt{3} - \Delta) - \hat{y}(\sqrt{3} - \Delta) = \Delta$, so $\hat{y}(\sqrt{3} - \Delta) = 1/(\sqrt{3} - \Delta) - \Delta$.

This construction does in fact lead to a stable matching. A firm with type $x \in [1, \sqrt{3}]$ matches either with a worker of type $1/x$ or with one of type

$$\hat{y}(x) \equiv \frac{1}{x} - \left(\sqrt{3} - x\right) = \frac{1}{x} + x - \sqrt{3}.$$

The need to match equal measures of firms and workers pins down the probability of each of a firm's two potential matches. If a firm of type x matches with a worker of type $1/x$ with probability $1/x^2$, then the total measure of firms in those matches equals the measure of available workers: for any positive $\Delta \leq \sqrt{3} - 1$,

$$\int_{\sqrt{3}-\Delta}^{\sqrt{3}} \frac{1}{x^2} \cdot \frac{1}{2} dx = \frac{1}{2}\left(\frac{1}{\sqrt{3} - \Delta} - \frac{1}{\sqrt{3}}\right) = \int_{\frac{1}{\sqrt{3}}}^{\frac{1}{\sqrt{3}-\Delta}} \frac{1}{2} dy.$$

With the complementary probability $1 - 1/x^2$, a firm of type x matches with a worker of type $\hat{y}(x)$.

It remains to verify that each firm of type $x \in [1, \sqrt{3})$ prefers high effort from a worker of type $\hat{y}(x)$ to no effort from worker type $y = \sqrt{3}$. Since

$$-1 + x\sqrt{3} < 1 + x\left(\frac{1}{x} + x - \sqrt{3}\right)$$

in that range, $\hat{y}(x)$ is in fact the firm's second choice after $y = 1/x$. Thus, in the proposed matching firms with types in $[1, \sqrt{3}]$ match with the workers with types in $[2 - \sqrt{3}, 1]$.

For the "low middle" region for firm types, $x \in \left[1/\sqrt{3}, 1\right]$, the opposite problem arises: the range of firms with such a type x is smaller than the range $[1, \sqrt{3}]$ of their preferred worker types $1/x$. After matching those firms, there will be a leftover mass of workers equal to $\frac{1}{2}(\sqrt{3} - 1) - \frac{1}{2}(1 - 1/\sqrt{3}) \approx 0.15$. Those workers then match with the next available firms, those with types from $x = 2 - \sqrt{3} \approx 0.27$ to $x = 1/\sqrt{3} \approx 0.58$, in positively assortative fashion. The matching probabilities mirror those in the high middle region, switching the roles of firms and workers: a worker of type $y \in [1, \sqrt{3}]$ matches with a firm of type $1/y$ with probability $1/y^2$, and with a firm of type $\frac{1}{y} + y - \sqrt{3}$ with probability $1 - 1/y^2$. The firms with types in $[2-\sqrt{3}, 1]$ match with the workers with types in $[1, \sqrt{3}]$.

The matching $\tilde{\mu}$ just constructed is summarized here:

$$\tilde{\mu}(x) = \begin{cases} x & \text{if } x \geq \sqrt{3} \\ \frac{1}{x} \text{ w/prob. } \frac{1}{x^2}, \hat{y}(x) \text{ w/prob. } 1 - \frac{1}{x^2} & \text{if } 1 \leq x < \sqrt{3} \\ \frac{1}{x} & \text{if } \frac{1}{\sqrt{3}} \leq x < 1 \\ \hat{y}^{-1}(x) & \text{if } 2 - \sqrt{3} \leq x < \frac{1}{\sqrt{3}} \\ x & \text{if } x < 2 - \sqrt{3} \end{cases}$$

Recall that $\hat{y}(x) \equiv \frac{1}{x} + x - \sqrt{3}$, so that for values of x between 1 and $\sqrt{3}$, $\hat{y}$ is strictly increasing and maps to $(2 - \sqrt{3}, 1/\sqrt{3})$. Thus, its inverse,

$$\hat{y}^{-1}(x) \equiv \frac{1}{2}\left(x + \sqrt{3} + (x^2 + 2x\sqrt{3} - 1)^{\frac{1}{2}}\right),$$

is well-defined. Figure 3.3 illustrates the matching $\tilde{\mu}$. The thickness of the matching curve represents the probability of the match.

3.2.1. *Effort levels.* Note, however, that the outcome $(\tilde{\mu}, \gamma^{\max})$, where the worker in each match exerts the highest effort level consistent with equilibrium, is not stable. The reason is that a firm inside the high middle region, with a type $x \in (1, \sqrt{3})$, strictly prefers one of his possible partners $(y = 1/x)$ to the other $(y = \hat{y}(x))$ if both exert full effort, as the the equilibrium selection rule $\gamma^{\max}$ specifies. But then a type-x firm matched with a type-$\hat{y}(x)$ worker and a type-$(1/x)$ worker matched with a different type-x firm form a blocking pair. If those two match with each other and the worker exerts effort e' just below 1, then both do strictly better than in their assigned match and equilibrium. To get stability, we adjust γ so that the firm is indifferent between its two possible matches. In particular, a firm of type x gets effort 1 if it is matched with a worker of type $\hat{y}(x)$, and effort $\hat{e}(x) \equiv 1 - \frac{1}{2}x(\sqrt{3} - x)$ if it is matched with a worker of type $1/x$; the firm's payoff is $2 - x(\sqrt{3} - x)$ in either case.

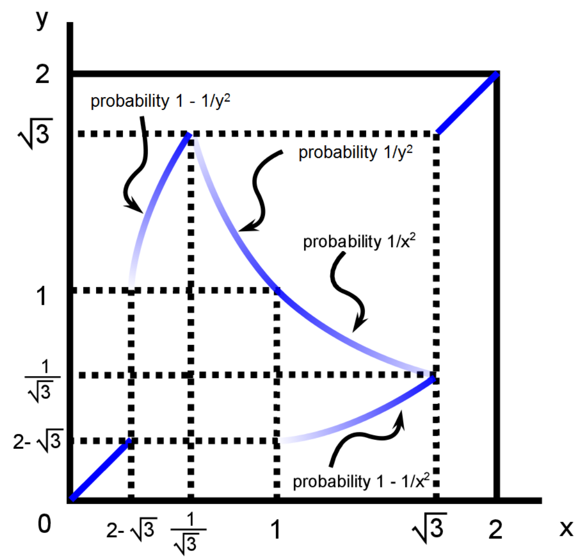

FIGURE 3.3.

In the low middle region of firms, where the firms are outnumbered by their preferred workers, specifying full effort does not cause a problem for stability. A worker with a type $y \in (1, \sqrt{3})$ who has matched with a firm of type $x = \hat{y}(y)$ would strictly prefer to match with his other possible partner ($x = 1/y$). But the firms of that type are already getting effort 1 from another worker of type y, so the worker cannot entice such a firm to switch.

One last issue involves the "left over" workers with types between 1 and $\sqrt{3}$: the match between a firm of type x' just above $2-\sqrt{3}$ and a worker of type y' just above 1 must specify effort strictly less than 1. Otherwise, the type-y' worker could offer effort just below 1 to a firm of type x just below $2-\sqrt{3}$ (matched with a worker of type $y = x < x'$ and getting effort 1) and form a blocking pair.[4] In particular, the type-y' worker exerts effort

$$\check{e}(y') \equiv \min\left\{1, \frac{13}{4} - 2\sqrt{3} - \frac{3}{2}\ln(y') + \frac{1}{4}(y')^2 + \frac{\sqrt{3}}{2}y'\right\}. \tag{3.1}$$

[4]In fact, because the function $\hat{y}^{-1}(x)$ increases so steeply for x just above $2-\sqrt{3}$, stability requires that effort is decreasing in firm type in that range. That requirement yields (3.1).

Thus, the matching $\tilde{\mu}$ is stable if the equilibrium level of effort $\tilde{e}$ in each match is as follows:

$$\tilde{e}(x,\tilde{\mu}(x)) = \begin{cases} 0 & \text{if } x \geq \sqrt{3} \\ \check{e}\left(\hat{y}^{-1}(x)\right) & \text{if } 2-\sqrt{3} \leq x < \frac{1}{\sqrt{3}} \\ \hat{e}(x) & \text{if } 1 \leq x < \sqrt{3} \text{ and } \tilde{\mu}(x) = \frac{1}{x} \\ 1 & \text{otherwise} \end{cases}$$

Call the resulting equilibrium payoffs $\tilde{\gamma}$:

$$\tilde{\gamma}(x,\tilde{\mu}(x)) = \begin{cases} (-1+x^2, 2+x^2) & \text{if } x \geq \sqrt{3} \\ \left(2-x(\sqrt{3}-x), 2+\frac{1}{2}x(\sqrt{3}-x)\right) & \text{if } 1 \leq x < \sqrt{3} \text{ and } \tilde{\mu}(x) = \frac{1}{x} \\ \left(2-x(\sqrt{3}-x), 2-x(\sqrt{3}-x)\right) & \text{if } 1 \leq x < \sqrt{3} \text{ and } \tilde{\mu}(x) = \hat{y}(x) \\ (2,2) & \text{if } \frac{1}{\sqrt{3}} \leq x < 1 \\ (-1+2\check{e}\left(\hat{y}^{-1}(x)\right)+x\hat{y}^{-1}(x), & \text{if } 2-\sqrt{3} \leq x < \frac{1}{\sqrt{3}} \\ \quad 2-\check{e}\left(\hat{y}^{-1}(x)\right)+x\hat{y}^{-1}(x)) & \\ (1+x^2, 1+x^2) & \text{if } x < 2-\sqrt{3} \end{cases}$$

Then the equilibrium selection rule $\tilde{\gamma}$ together with the matching $\tilde{\mu}$ represents a stable outcome in the continuous limit, in the following sense: there exists a sequence of sets of agent types $(\mathcal{X}_N, \mathcal{Y}_N)_N$ such that 1) the type distributions for both firms and workers converge weakly to the continuous uniform distribution over $\left[\underline{\theta}, \bar{\theta}\right]$, and 2) for every N there is a (deterministic) matching μ_N satisfying $\mu_N(x) \in \text{supp}\tilde{\mu}(x)$ for all $x \in \mathcal{X}_N$ such that the outcome $(\mu_N, \tilde{\gamma})$ is stable for every N. The appendix gives a proof of that claim.

A firm's payoff as a function of its type x is graphed in Figure 3.4. Note in particular that the payoff is not monotonic in the firm's type: if x is just below 1, then the firm gets a payoff of 2, while a firm with a type just above 1 gets payoff $3-\sqrt{3} \approx 1.27$.

Figure 3.5, similarly, shows a worker's payoff as a function of its type y. A worker with a type y between 1 and $\sqrt{3}$ gets one of two possible payoffs, depending on whether he matches with a firm of type $x = 1/y$ or with a firm of type $\hat{y}(y)$. The first match occurs with probability $1/y^2$ and yields payoff 2; the second occurs with probability $1-1/y^2$ and yields payoff $2-\check{e}\,(y)+y\cdot\hat{y}(y)$. The dotted line in Figure

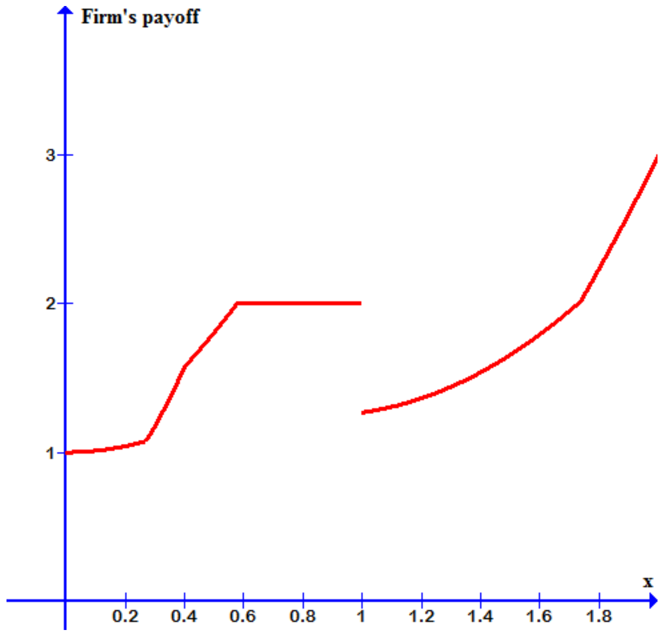

FIGURE 3.4. Firms' payoffs under the stable outcome $(\tilde{\mu}, \tilde{\gamma})$

3.5 shows the expected payoff. As with the firms, a worker's payoff is higher if his type is just below 1 than if it is just above. Intuitively, these nonmonotonicities arise from the fact that the function mapping a firm's type x to the type of the worker that it is just willing to fire, $1/x$, is not measure-preserving. For $x > 1$, the slope of that function is less than 1 in absolute value, so workers are the "short side" of the market. As described above, the resulting competition among relatively many firms for relatively few workers pushes down the firms' payoffs. For $x < 1$, the absolute value of the slope is greater than 1, and it is the firms who are in short supply.

Crawford (1991) and Ashlagi *et al.* (forthcoming) consider the effects on stable matching and payoffs of having more agents on one side of the market than on the other. Here, the imbalances arise endogenously in a setting where the two sides of the market are symmetric.

The stable outcome $(\tilde{\mu}, \tilde{\gamma})$ constructed above for the continuous limit (as N grows without bound) features "fractional matching" (Roth *et al.*, 1993): an agent of a given type matches with two distinct types of partner with positive probability. For a large but fixed N, each agent type matches with a single partner type, but agents whose

FIGURE 3.5. Workers' payoffs under the stable outcome $(\tilde{\mu}, \tilde{\gamma})$

types are very close may match with very different partner types. Figure 3.6 shows an example.

4. General Results

In this section, we show how the results from the example generalize. A key step in the argument that positively assortative matching is not stable in the example is the following: if a firm of type x is matched with a worker of type x, and a type-$(x+\epsilon)$ firm with a worker of type $x+\epsilon$, and effort is 0 in both matches, then there is an effort level $e' > 0$ such that both the type-$(x+\epsilon)$ firm and the type-x worker benefit from matching with each other at effort e'. That is, at low effort levels firms are willing to trade quality for effort at a higher rate than are workers. We generalize that condition as follows:

Definition (Firms care more about effort)**.** *Firms care more about effort than workers do* if for any type $\theta \in \left[\underline{\theta}, \bar{\theta}\right]$ and any effort level $e \in [0,1]$,

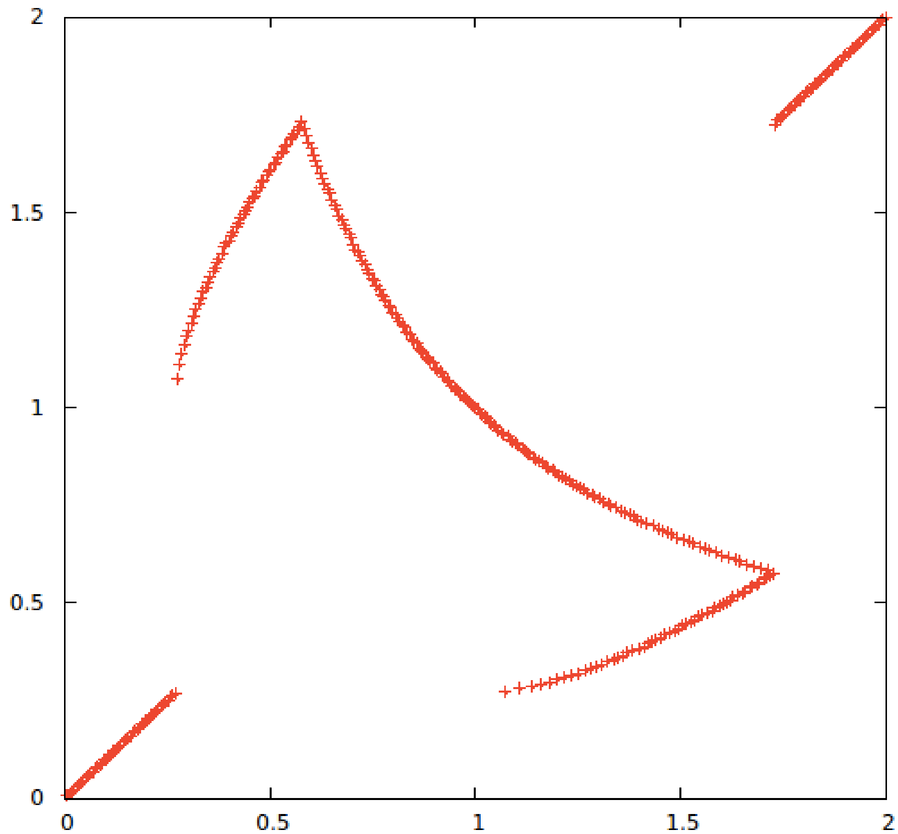

FIGURE 3.6. A discrete version of the stable matching $\tilde{\mu}$

$$\frac{U_e^F(0;\theta,\theta)}{U_y^F(0;\theta,\theta)} > -\frac{U_e^W(0;\theta,\theta)}{U_x^W(0;\theta,\theta)}. \tag{4.1}$$

If, all else equal, the marginal benefit of low effort is large enough or the marginal cost is small enough, then firms will care more about effort than workers do.

Let $(\mathcal{X}_N, \mathcal{Y}_N)_N$ be a sequence of sets of agent types such that the type distributions for both firms and workers converge weakly to the continuous uniform distribution over $\left[\underline{\theta}, \bar{\theta}\right]$.

Theorem 1. *If firms care more about effort than workers do, then for large enough N positively assortative matching is not stable.*

The argument for why positively assortative matching is unstable if firms care strictly more about effort is basically identical to the argument for the example in Section 3.1. Kaneko (1982) establishes the existence of a stable matching in this environment. That finding, together with Theorem 1, immediately implies the following result:

Corollary 1. *If firms care more about effort than workers do, then for large enough N there exists a stable matching that is not positively assortative.*

Thus, as long as firms care more about effort than workers do, the results from the example for the case without wages generalize.

4.1. **The structure of the set of stable outcomes.** Recall that under either NTU or TU matching, the set of stable outcomes has a lattice structure: if we take any two stable outcomes, then a third outcome where each agent on one side of the market gets his preferred assignment from among the original two outcomes is itself stable. Further, that third outcome also gives each agent on the other side of the market his less preferred assignment. It follows that there exists a stable outcome that is optimal (among the set of all stable outcomes) for every agent on one side of the market, and that that same outcome is the least preferred stable outcome for all the agents on the other side.

A version of that result also holds in our setting. In the case of non-transferable utility, the lattice result requires that all agents have strict preferences over their partners. With transferable utility, on the other hand, Demange and Gale's (1985) result allows indifference. The reason for that contrast, briefly, is that for any outcome (μ, γ), if there exists a pair (x, y) and a transfer between them such that the type-y worker strictly prefers that outcome to (μ, γ) and the type-x firm weakly prefers it, then (μ, γ) is not stable: the worker could transfer an additional $\epsilon > 0$ of utility to the firm with the result that now both firm and worker strictly prefer to switch. In our setting, that argument breaks down because there is a bound on how much utility can be transferred within a match. If, for example, effort in a match is already equal to 1, then the worker cannot transfer any additional utility to the firm.

To derive the lattice result in our setting, therefore, we adapt Demange and Gale's (1985) proof, but we need to rule out the problems that ties in preferences introduce in the NTU case and that persist with imperfectly transferable utility. Along those lines, we introduce the following definition. We say that an outcome is *strictly stable* if no worker and firm can match with each other and achieve an equilibrium payoff that both of them weakly prefer to their current outcome and that at least one of them strictly prefers. In an abuse of notation, for any outcome (μ, γ), we will write $\gamma(x)$ and $\gamma(y)$ as the payoffs that a type-x firm and a type-y worker receive, respectively, under (μ, γ). That is, $\gamma(x) = \gamma(x, \mu(x))$ and $\gamma(y) = \gamma(\mu^{-1}(y), y)$.

Definition (Strict stability)**.** An outcome (μ, γ) is *strictly stable* if there do not exist a firm of type x, a worker of type y, and a payoff $\left(\hat{\gamma}^F, \hat{\gamma}^W\right) \in E^\delta(x, y)$ such that $\hat{\gamma}^F \geq \gamma(x)$ and $\hat{\gamma}^W \geq \gamma(y)$, with at least one strict inequality.

Strict stability is a more demanding notion than stability, because an outcome fails to be strictly stable if some pair can generate a weak (rather than a strict) Pareto improvement for themselves.

Using that definition, we can state the lattice result.

Theorem 2. *Let (μ^A, γ^A) and (μ^B, γ^B) be strictly stable outcomes. Then there exist strictly stable outcomes $(\bar{\mu}, \bar{\gamma})$ and $(\underline{\mu}, \underline{\gamma})$ such that for all $x \in \mathcal{X}_N$ and $y \in \mathcal{Y}_N$,*

- $\bar{\gamma}(x) = \max\{\gamma^A(x), \gamma^B(x)\}$ *and* $\bar{\gamma}(y) = \min\{\gamma^A(y), \gamma^B(y)\}$*, and*
- $\underline{\gamma}(x) = \min\{\gamma^A(x), \gamma^B(x)\}$ *and* $\underline{\gamma}(x) = \max\{\gamma^A(x), \gamma^B(x)\}$*.*

5. Efficiency Wages

So far we have studied the patterns of matchings that emerge when the wage paid by the firm is fixed at an exogenous market rate (which may, however, vary with firm and worker types). That market wage is already incorporated into the payoff functions $U^W(\cdot)$ and $U^F(\cdot)$. In this section, we allow firms to make additional payments beyond the market wage. We can then use our model to examine, for a given matching, how a firm can use wages to incentivize effort. In particular, we ask when a firm does better by committing to a flat wage that does not condition on effort than by promising a bonus after high effort. We interpret the former as an efficiency wage.

Suppose that a firm of type x is matched with a worker of type y, and that $U^F(0;x,y) > \underline{U}^F(x)$, so that at the market wage effort is not sustainable, and the firm gets payoff $U^F(0;x,y)$ in equilibrium. Even though Condition 2.1 implies that the threat of firing is enough to motivate effort by the worker, the firm is unwilling to fire because its payoff even from zero effort exceeds its outside option. The firm has two ways to use additional wages to get the worker to exert full effort. First, it can pay the worker a bonus $b > 0$ in each period in which $e = 1$. The optimal bonus for the firm, $b(x,y)$, just equals the worker's cost of effort: $b(x,y) = U^W(0;x,y) - U^W(1;x,y)$. The resulting payoff, $V^b(x,y)$, for the firm is

$$V^b(x,y) \equiv U^F(1;x,y) - b(x,y) = U^F(1;x,y) - \left[U^W(0;x,y) - U^W(1;x,y)\right].$$

Second, suppose that the firm can commit to paying an additional per-period wage $w > 0$ for as long as the worker is employed – that is, as long as the firm does not fire him. Such a wage can affect effort incentives indirectly: it lowers the firm's payoff for any effort level and so increases its willingness to fire the worker. The credible threat of firing can then enforce effort in equilibrium, since Condition 2.1 holds. The optimal additional wage for the firm, $w(x,y)$, is the one that makes it just indifferent between 1) paying $w(x,y)$ plus the market wage and getting zero effort, and 2) receiving its outside option: $w(x,y) = U^F(0;x,y) - \underline{U}^F(x)$. The resulting payoff for the firm, $V^w(x,y)$, is

$$V^w(x,y) \equiv U^F(1;x,y) - w(x,y) = U^F(1;x,y) - \left[U^F(0;x,y) - \underline{U}^F(x)\right].$$

We interpret this case as an efficiency wage, a payment level that is not based on performance and that is above the market rate.

Krueger and Summers (1988, p.261) identify four theories to explain efficiency wages: 1) to reduce turnover, 2) to increase effort by raising the worker's cost of losing the job, 3) to create feelings of loyalty toward the firm and thus increase workers' productivity, and 4) to attract higher-quality workers. The explanation here, that a higher wage indirectly improves effort incentives by making the threat of firing credible, is distinct from all four of those rationales.

Which firms will choose to pay efficiency wages? That is, when is $V^w(x,y)$ greater than both $V^b(x,y)$ and $U^F(0;x,y)$, given that $U^F(0;x,y) > \underline{U}^F(x)$? First, the firm prefers to pay the efficiency wage $w(x,y)$ rather than the bonus $b(x,y)$ when $w(x,y) < b(x,y)$. That inequality holds when

$$U^F(0;x,y) < \underline{U}^F(x) + \left[U^W(0;x,y) - U^W(1;x,y)\right].$$

Second, the firm prefers paying $w(x,y)$ and getting effort 1 to getting effort 0 when

$$U^F(0;x,y) < \frac{1}{2}\left[U^F(1;x,y) + \underline{U}^F(x)\right].$$

Overall, the firm prefers paying the efficiency wage when

$$U^F(0;x,y) < \min\left\{\frac{1}{2}\left[U^F(1;x,y) + \underline{U}^F(x)\right], \underline{U}^F(x) + \left[U^W(0;x,y) - U^W(1;x,y)\right]\right\}.$$

Intuitively, efficiency wages are attractive when the cost of effort for the worker is high (so that a performance-based bonus is costly), and when the value of effort to the firm is also high (so that accepting low effort is unappealing). The condition that $U^F(0;x,y) > \underline{U}^F(x)$, so that without an additional wage the firm is unwilling to fire, holds when even minimal output from the worker is still better than the firm's outside option. That condition would hold, for example, when the worker is highly skilled and difficult to replace. To summarize, our analysis predicts that efficiency wages will be present in industries where workers are productive even without high effort, and where effort is very costly for workers and very valuable to the firm.[5]

[5]The empirical literature on efficiency wages gives some support to those predictions. Campbell (1993, p.464) finds that, after controlling for worker characteristics, "the wage depends positively on firm characteristics that raise workers' productivity." Krueger and Summers (1988) use industry dummy variables rather than measures of firm characteristics that affect worker producitivity. They estimate, for example, that "the average employee in the mining industry earns wages that are 24 per cent higher than the average employee in all industries, after controlling for human capital and demographic background" (p.264). That finding is consistent with our analysis, if it is plausible to think that mining fits the criteria of high effort cost, high value of effort, and high value of output even with low effort. Other industries at the two-digit census industry code level with above-average wages in Krueger and Summers' (1988) study include chemicals and petroleum, while private household servies, education services, and welfare services pay wages below the average.

6. Discussion

In this paper, we endogenize the value of a match between a firm and a worker as the outcome of a dynamic game: in each period the worker chooses his effort level and the firm decides whether or not to continue the relationship. We find that positively assortative matching may not be stable even though payoffs are increasing and supermodular in types. We identify a new interpretation of efficiency wages: by paying workers above the level required to induce effort, firms increase their willingness to fire high-quality workers and thus indirectly improve effort incentives.

In the context of a parametrized example, we construct a stable outcome that features positively assortative matching at the top and bottom of the type distribution, and negatively assortative matching in the middle. Although the distributions of firms and workers are symmetric, the nonlinear negatively assortative matching in the middle generates endogenous imbalances. As a result, the matching is extremely discontinuous.

Our model specifies the continuation value to an agent after a firing as an exogenous function of the agent's type. Our results rely on a single-crossing condition: the value of that outside option for the firm increases in type more slowly than the value of the stage-game payoff. That assumption would be satisfied if, for example, we extended the model to allow re-matching, and the outside option were the match value multiplied by (1 minus the expected discounted time to find a new partner). In that case, the outside option is a fixed fraction of the stage-game payoff, so the single-crossing condition holds. Endogenizing the post-match continuation value within a model of matching with search frictions is a subject for future research.

Appendix A. Proofs

A.1. **Proving Lemma 1.**

Proof. If $U^F(0;x,y) > \underline{U}^F(x)$, then for any strategy of the worker, the firm's payoff from firing is strictly lower than its payoff from continuing. Thus, in any SPE the firm will continue after every history. The worker's best response is to always choose effort 0.

If $U^F(0;x,y) \leq \underline{U}^F(x)$, then the strategy profile $\underline{\sigma}$ where after any history the firm chooses to fire and the worker chooses effort 0 is an SPE; $\underline{\sigma}$ yields payoffs $(\underline{U}^F(x), \underline{U}^W(y))$. For any effort level $\hat{e} \geq \underline{e}(x,y)$, Condition 2.1 ensures that the threat of reversion to $\underline{\sigma}$ supports the outcome where the firm does not fire and the worker chooses effort level $\hat{e}$ in each period. Thus, $\text{co}\left\{\hat{V}(x,y), \left(\underline{U}^F(x), \underline{U}^W(y)\right)\right\} \subseteq E^\delta(x,y)$. Any feasible payoff outside $\text{co}\left\{\hat{V}(x,y), \left(\underline{U}^F(x), \underline{U}^W(y)\right)\right\}$ gives the firm less than its minmax payoff $\underline{U}^F(x)$ and so cannot lie in $E^\delta(x,y)$. We conclude that $E^\delta(x,y) = \text{co}\left\{\hat{V}(x,y), \left(\underline{U}^F(x), \underline{U}^W(y)\right)\right\}$. □

A.2. **Proving that $(\tilde{\mu}, \tilde{\gamma})$ is stable in the example.**

Proof. We will show that in the continuous limit, under $(\tilde{\mu}, \tilde{\gamma})$ there are no blocking pairs. It follows immediately that for any $(\mathcal{X}_N, \mathcal{Y}_N)$ and matching μ_N such that $\mu_N(x) \in \text{supp}\tilde{\mu}(x)$ and $\mu_N(x) \in \mathcal{Y}_N$ for every $x \in \mathcal{X}_N$, $(\mu_N, \tilde{\gamma})$ is stable with respect to $(\mathcal{X}_N, \mathcal{Y}_N)$.

We construct the desired sequence $(\mathcal{X}_N, \mathcal{Y}_N)_N$ as follows. First, choose any sequence $(\mathcal{X}_N)_N$ such that the distribution of firm types converges weakly to the continuous uniform distribution over $\left[\underline{\theta}, \bar{\theta}\right]$. For each N, we then generate a random set $\mathcal{Y}_N$. Let

$$\mathcal{X}_N^2 \equiv \mathcal{X}_N \cap \left[1, \sqrt{3}\right)$$

represent the set of firm types that under $\tilde{\mu}$ may match with two different types of worker. For each $x_n \in \mathcal{X}_N^2$, define $y(x_n, N)$ as the realization of a random variable that takes on each of the two possible values of $\tilde{\mu}(x_n)$ with the specified probability. (That is, $y(x_n,N) = 1/x_n$ with probability $1/(x_n)^2$, and $y(x_n,N) = \hat{y}(x_n)$ with probability $1 - 1/(x_n)^2$.) The distributions of the $y(x_n,N)$'s are independent across both n and N.

The other firm types in $\mathcal{X}_N$, those lying in

$$\mathcal{X}_N^1 \equiv \mathcal{X}_N \cap \left[0, \sqrt{3}\right) \cup \left[\frac{1}{\sqrt{3}}, 1\right],$$

match with a unique worker type under $\tilde{\mu}$. We then define the elements of the sequence $\mathcal{Y}_N$ as

$$\left\{\cup_{x_n \in \mathcal{X}_N^1} \tilde{\mu}(x_n)\right\} \cup \left\{\cup_{x_n \in \mathcal{X}_N^2} y(x_n, N)\right\}.$$

That is, we construct $\mathcal{Y}_N$ to contain the match $\tilde{\mu}(x_n)$ of every $x_n \in \mathcal{X}_N$. For firm types $x_n \in \mathcal{X}_N$ that $\tilde{\mu}$ assigns randomly, the matching worker type in $\mathcal{Y}_N$ is also chosen randomly, with the same probabilities. Thus, with probability 1 the distributions of types in the sequence $(\mathcal{Y}_N)_N$ converges weakly to the continuous uniform distribution over $\left[\underline{\theta}, \bar{\theta}\right]$, as desired.

To show that there are no blocking pairs under $(\tilde{\mu}, \tilde{\gamma})$ in the continuous limit, we first observe that for every matching, $\tilde{\gamma}$ specifies a payoff on the Pareto frontier of the SPE set. Thus, no firm and worker that are matched with each other under $\tilde{\mu}$ can form a blocking pair.

Next, we note that for any matched pairs (x, y) and (x', y') with effort levels e and e', respectively, the firm of type x and the worker of type y' cannot form a blocking pair if

$$y' \cdot (x' - x) - \frac{1}{2} x \left[y' - y\right] \geq e' - e. \tag{A.1}$$

The reason is that they form a blocking pair only if there is an effort level $\hat{e} \in [0, 1]$ such that

$$-1 + 2\hat{e} + xy' > -1 + 2e + xy$$

and

$$2 - \hat{e} + xy' > 2 - e' + x'y'.$$

If Condition A.1 holds, then no such $\hat{e}$ exists.

Case 1. $x, y \in \left[\sqrt{3}, 1\right]$. Each firm with a type x in this region can get a higher payoff only by switching to a worker with a higher type, but each worker in this region also would only switch to a firm with a higher type. Thus, no firm or worker with a type in this region can be involved in any blocking pair, and they can be ignored for the rest of the analysis.

Case 2. $x \in [1, \sqrt{3}]$. By construction, each firm with a type x in this region is indifferent between its two possible matches, $1/x$ and $\hat{y}(x)$. The firm cannot form a blocking pair with any worker on the $\hat{y}$ branch (that is, $y \in [2-\sqrt{3}, 1/\sqrt{3}]$): the firm would switch only to a higher y in that range (since it is getting effort 1 from the worker of type $\hat{y}(x)$), and for any $x' > x$ in that range straightforward algebra gives

$$\begin{aligned} \hat{y}(x')(x'-x) - \tfrac{1}{2}x\left[\hat{y}(x') - \hat{y}(x)\right] &= \tfrac{(x'-x)}{x'}\tfrac{1}{2}\left((x')^2 2 - \sqrt{3}x' + 3\right) \\ &\geq 0 \\ &= 1 - 1 = e(x') - e(x). \end{aligned}$$

Thus, Condition A.1 holds, so no such blocking pair exists.

The firm also cannot form a blocking pair with any worker whose type y is below $2-\sqrt{3}$, again because the firm is already getting 1 from a worker of type $\hat{y}(x) \geq 2-\sqrt{3}$. For $y \in \left(1, \sqrt{3}\right)$, effort is not sustainable when the firm matches with a type-y worker, and simple algebra shows that the firm prefers effort 1 from a type-$\hat{y}(x)$ worker to effort 0 from a worker of type $\sqrt{3}$ (the best y in that range). Finally, for $y \in \left[1/\sqrt{3}, 1\right]$, similar reasoning means that we need only check for blocking pairs where the firm matches with a worker whose type is lower than $1/x$: matching with a higher type would lead to zero effort. For any $y < 1/x$ in this range, more algebra shows that Condition A.1 holds: there is no effort level $\hat{e}$ such that both the firm and the type-y worker prefer to match with each other at $\hat{e}$ over what they're getting currently.

Case 3. $x \in \left[1/\sqrt{3}, 1\right]$. Again, we need only check for blocking pairs where a firm with a type x in this range matches with a worker whose type y is lower than $1/x$ – effort is not sustainable when the firm is matched with a worker whose type is above that level, and the firm prefers effort 1 from its current partner (type $1/x$) to effort 0 from a worker of type $\sqrt{3}$ (the highest y potentially available). The firm would only prefer a worker with a lower y only if it could get higher effort, but it is already getting effort 1. Thus, firms with a type in this range are not part of any blocking pairs.

Case 4. $x \in \left[2-\sqrt{3}, 1/\sqrt{3}\right]$. A firm with a type x in this range gets a higher payoff from its current outcome than from getting effort 1 when matched with a worker of type $y < 2-\sqrt{3}$. Thus, it cannot form a blocking pair with such a worker. If $\check{e}\left(\hat{y}^{-1}(x)\right) < 1$ (that is, if x is below a cutoff value $\bar{x} \approx 0.4$), then for each worker type $y \in \left[2-\sqrt{3}, \sqrt{3}\right]$, straightforward but somewhat tedious algebra establishes that Condition A.1 holds, so the firm cannot be part of a blocking pair. If $\check{e}\left(\hat{y}^{-1}(x)\right) = 1$ (that is, if $x \geq \bar{x}$), then the same algebra shows that Condition A.1 holds for $y \in (\tilde{\mu}(x), \sqrt{3}]$. The firm also cannot be part of a blocking pair with a worker of type $y < \tilde{\mu}(x)$, because it gets effort 1 in its current match.

Case 5. $x \in \left[0, 2-\sqrt{3}\right]$. A firm with a type x in this range is matched with a worker of the same type and gets effort 1, so it cannot form a blocking pair with a worker of type $y < x$. For each worker type $y \in \left[2-\sqrt{3}, \sqrt{3}\right]$, more algebra establishes that Condition A.1 holds, so the firm cannot be part of a blocking pair.

□

A.3. **Proving Theorem 1.**

Proof. Under the positively assortative matching μ^+, positive effort is sustainable in equilibrium when $U^F(0; x_n, y_n) \leq \underline{U}^F(x)$. As N increases, the limit of μ^+ is the identity matching μ^I, where $\mu^I(x) = x$. Under μ^I, the condition for effort to be sustainable is $U^F(0; x, x) \leq \underline{U}^F(x)$. Recall from Assumption 3 that $\hat{\theta} \in \left(\underline{\theta}, \bar{\theta}\right)$ is the type such that $U^F(0; x, x) < \underline{U}^F(x)$ for $x < \hat{\theta}$, and $U^F(0; x, x) > \underline{U}^F(x)$ for $x > \hat{\theta}$.

Because $U^F(0; \hat{\theta}, \hat{\theta}) = \underline{U}^F(\hat{\theta})$, for any sequence of strictly positive ϵ_k converging to 0, we can choose $x(k), y(k)$ such that $x(k) \in \left(\hat{\theta}, \hat{\theta}+\epsilon_k\right)$, $y(k) \in \left(\hat{\theta}-\epsilon_k, \hat{\theta}\right)$, and $U^F(0; x(k), y(k)) < \underline{U}^F(x(k))$. We will show that under μ^I, in the continuous limit $x(k)$ and $y(k)$ form a blocking pair for large enough k. Then we will argue that for large N, there must be a similar blocking pair under μ^+.

To see that $x(k)$ and $y(k)$ are a blocking pair, note that since $U^F(0; x(k), y(k)) > \underline{U}^F(x(k))$, the type-$x(k)$ firm's match partner will not exert effort in equilibrium, and so the firm's payoff will be $U^F(0; x(k), y(k))$. The firm would prefer to match with the type-$y(k)$ worker if that worker exerted effort e^* strictly above $\underline{e}(k)$, defined implicitly as

$$U^F(\underline{e}(k); x(k), y(k)) = U^F(0; x(k), x(k)).$$

The type-$y(k)$ worker's current payoff is at most $U^F(0; y(k), y(k))$, so he also prefers to match with the type-$x(k)$ firm and exert effort e^* as long as e^* is strictly below $\bar{e}(k)$, defined by

$$U^F(\bar{e}(k); w(k), y(k)) = U^F(0; y(k), y(k)).$$

Taking a first-order approximation, for large k, we get

$$\underline{e}(k) \approx (x(k) - y(k)) \frac{U_y^F(0; \hat{\theta}, \hat{\theta})}{U_e^F(0; \hat{\theta}, \hat{\theta})}$$

and

$$\bar{e}(k) \approx (x(k) - y(k)) \frac{U_x^W(0; \hat{\theta}, \hat{\theta})}{-U_e^W(0; \hat{\theta}, \hat{\theta})}.$$

Since both $x(k)$ and $y(k)$ converge to $\hat{\theta}$, $\underline{e}(k) < 1$ for large k. Taking the ratio, we get that for large enough k,

$$\frac{\bar{e}(k)}{\underline{e}(k)} \approx \frac{\frac{U_x^W(0;\hat{\theta},\hat{\theta})}{-U_e^W(0;\hat{\theta},\hat{\theta})}}{\frac{U_y^F(0;\hat{\theta},\hat{\theta})}{U_e^F(0;\hat{\theta},\hat{\theta})}}.$$

Condition 4.1 then implies that $\bar{e}(k) > \underline{e}(k)$ for large k. Because $U^F(0; x(k), y(k)) < \underline{U}^F(x(k))$, any effort level $e^* \in [\underline{e}(k), \bar{e}(k)]$ is sustainable in equilibrium when the type-$x(k)$ firm and the type-$y(k)$ worker are matched. Thus, they form a blocking pair.

For a given N and k, let $x_{n^*(N,k)}$ be the closest type of firm in $\mathcal{X}_N$ to $x(k)$, and let $y_{n^{**}(N,k)}$ be the closest type of worker in $\mathcal{Y}_N$ to $y(k)$. As N increases, $x_{n^*(N,k)}$ and $y_{n^*(N,k)}$ converge to $x(k)$, and $y_{n^{**}(N,k)}$ and $x_{n^{**}(N,k)}$ converge to $y(k)$. Therefore, when k and N are large enough, under μ^+ $x_{n^*(N,k)}$ and $y_{n^{**}(N,k)}$ form a blocking pair, and positively assortative matching is not stable. □

A.4. **Proving Theorem 2.** The first step is to derive the following analog of Demange and Gale's (1985) Lemma 1. Given two outcomes (μ^A, γ^A) and (μ^B, γ^B), define $X^{A \succ B}$, $X^{B \succ A}$, and $X^{A \sim B}$ as the sets of firm types who, respectively, get a strictly higher payoff under (μ^A, γ^A), get a strictly higher payoff under (μ^B, γ^B), and get the same payoff under both outcomes. Let $X^{A=B} \subseteq X^{A \sim B}$ denote the types of firms who 1) match with the same worker under μ^A and μ^B (that is, $\mu^A(x) = \mu^B(x)$), and 2) get the same payoff under both outcomes. Define $Y^{A \succ B}$, $Y^{B \succ A}$, $Y^{A \sim B}$ and $Y^{A=B}$ analogously.

Lemma 2. *Let two stable outcomes (μ^A, γ^A) and (μ^B, γ^B) be given. Suppose that no agent gets the same payoff in both outcomes unless he matches with the same partner under μ^A and μ^B; i.e. $X^{A \sim B} \setminus X^{A=B}$ and $Y^{A \sim B} \setminus Y^{A=B}$ are empty. Then both μ^A and μ^B match $X^{A \succ B}$ with $Y^{B \succ A}$ and $X^{B \succ A}$ with $Y^{A \succ B}$.*

Proof. The proof is by contradiction. Suppose that μ^A matches a type-x firm with a type-y worker, where $(x, y) \in X^{A \succ B} \times Y^{A \succ B}$. Then since both the type-x firm and the type-y worker get a strictly higher payoff under (μ^A, γ^A) than under (μ^B, γ^B), they form a blocking pair under (μ^B, γ^B), contradicting the stability of (μ^B, γ^B). Thus, under μ^A every firm with a type in $X^{A \succ B}$ must match with a worker whose type is in $Y^{B \succ A}$, and every worker with a type in $Y^{A \succ B}$ must match with a firm whose type is in $X^{B \succ A}$.

A symmetric argument shows that under μ^B every agent who prefers (μ^B, γ^B) must match with an agent who prefers (μ^A, γ^A). The result follows. □

The next lemma shows that for any two strictly stable outcomes, there is a nearby pair of strictly stable outcomes with the same matchings where every agent has a strict preference across the two outcomes.

Lemma 3. *Let two strictly stable outcomes (μ^A, γ^A) and (μ^B, γ^B) be given. Then for any $\epsilon > 0$, there exist strictly stable outcomes $(\mu^A, \tilde{\gamma}^A)$ and $(\mu^B, \tilde{\gamma}^B)$ such that*

(1) $X^{\tilde{A} \sim \tilde{B}} \setminus X^{\tilde{A}=\tilde{B}}$ and $Y^{\tilde{A} \sim \tilde{B}} \setminus Y^{\tilde{A}=\tilde{B}}$ are both empty, and
(2) for all $x \in \mathcal{X}_N$,

$$\max\left\{\left\|\tilde{\gamma}^A(x, \mu^A(x)) - \gamma^A(x, \mu^A(x))\right\|, \left\|\tilde{\gamma}^B(x, \mu^B(x)) - \gamma^B(x, \mu^B(x))\right\|\right\} < \epsilon.$$

Proof. The definition of strict stability implies that if (μ, γ) is a strictly stable outcome, then there exists $\eta > 0$ such that the following holds: for any equilibrium selection rule γ' such that for each x, 1) $\gamma'(x, \mu(x))$ is within η of $\gamma(x, \mu(x))$, and 2) $\gamma'(x, \mu(x))$ specifies a payoff on the Pareto frontier of the SPE set $E^\delta(x, \mu(x))$, the outcome (μ, γ') is also strictly stable. Any switch from a strictly stable outcome (μ, γ) – either two agents leaving their partners and matching together or two already-matched agents changing their equilibrium play – makes at least one member of the potential blocking pair strictly worse off; thus, a small change in γ preserves that feature as long as the change is not a Pareto worsening.

If $X^{A\sim B} \setminus X^{A=B}$ and $Y^{A\sim B} \setminus Y^{A=B}$ are empty, then the lemma holds trivially. Suppose that there exists a type of firm $x \in X^{A\sim B} \setminus X^{A=B}$. From the argument above, if we move either $\gamma^A(x, \mu^A(x))$ or $\gamma^B(x, \mu^B(x))$ slightly along the Pareto frontier of $E^\delta(x, \mu^A(x))$ or $E^\delta(x, \mu^B(x))$, then we break the type-x firm's indifference while preserving strict stability. Futher, if the change is small enough, then it will not change the preferences over outcomes of the type-x firm's partners $\mu^A(x)$ or $\mu^B(x)$ if they originally had a strict preference. If the partners were themselves indifferent, then the change will also break their indifference. By Lemma 1, such a change is possible unless both SPE sets $E^\delta(x, \mu^A(x))$ and $E^\delta(x, \mu^B(x))$ are singletons – that is, unless

$$E^\delta(x, \mu^i(x)) = \left\{U(0; x, \mu^i(x))\right\}, i \in \{A, B\}. \tag{A.2}$$

Since $x \in X^{A\sim B} \setminus X^{A=B}$, Condition A.2 cannot hold. Suppose that it did. The type-x firm is indifferent between the two outcomes, so it must be that $U^F(0; x, \mu^A(x)) = U^F(0; x, \mu^B(x))$. Because U^F is strictly increasing in worker type, that equality implies that $\mu^A(x) = \mu^B(x)$, contradicting the assumption that $x \notin X^{A=B}$. Thus, the desired change in either $\gamma^A(x, \mu^A(x))$ or $\gamma^B(x, \mu^B(x))$ is possible.

A symmetric argument applies for any worker type $y \in Y^{A\sim B} \setminus Y^{A=B}$, so we can construct the specified $\tilde{\gamma}^A$ and $\tilde{\gamma}^B$. □

Now we can complete the proof of Theorem 2.

Proof of Theorem 2. First, we construct a strictly stable outcome $(\bar{\mu}, \bar{\gamma})$ such that $\bar{\gamma}(x) = \max\{\gamma^A(x), \gamma^B(x)\}$ and $\bar{\gamma}(y) = \min\{\gamma^A(y), \gamma^B(y)\}$ for all $x \in \mathcal{X}_N$ and $y \in \mathcal{Y}_N$.

Let $(\mu^A, \tilde{\gamma}^A)$ and $(\mu^B, \tilde{\gamma}^B)$ be the strictly stable outcomes from Lemma 3, and define the matching $\bar{\mu}$ as follows:

$$\bar{\mu}(x) = \begin{cases} \mu^A(x) & \text{if } x \in X^{\tilde{A} \succ \tilde{B}} \\ \mu^B(x) & \text{otherwise.} \end{cases}$$

Since $X^{\tilde{A} \sim \tilde{B}} \setminus X^{\tilde{A} = \tilde{B}}$ and $Y^{\tilde{A} \sim \tilde{B}} \setminus Y^{\tilde{A} = \tilde{B}}$ are both empty, Lemma 2 ensures that $\bar{\mu}$ is a properly defined matching: the image of $X^{\tilde{A} \succ \tilde{B}}$ is the same under both μ^A and μ^B. Define the equilibrium selection rule $\bar{\gamma}$ similarly:

$$\bar{\gamma}(x, \bar{\mu}(x)) = \begin{cases} \gamma^A(x, \mu^A(x)) & \text{if } x \in X^{\tilde{A} \succ \tilde{B}} \\ \gamma^B(x, \mu^B(x)) & \text{otherwise.} \end{cases}$$

That equilibrium selection rule yields the desired payoff for each agent.

To see that $(\bar{\mu}, \bar{\gamma})$ is strictly stable, suppose conversely that there exist a firm of type x, a worker of type y, and a payoff $\left(\hat{\gamma}^F, \hat{\gamma}^W\right) \in E^\delta(x, y)$ such that

$$\hat{\gamma}^F \geq \bar{\gamma}(x) = \max\{\gamma^A(x), \gamma^B(x)\}$$

and

$$\hat{\gamma}^W \geq \bar{\gamma}(y) = \min\{\gamma^A(y), \gamma^B(y)\},$$

with at least one strict inequality. If $y \in X^{\tilde{B} \succ \tilde{A}}$, then $\bar{\gamma}(y) = \gamma^A(y)$. But then x, y, and $\left(\hat{\gamma}^F, \hat{\gamma}^W\right)$ contradict the strict stability of (μ^A, γ^A). Otherwise, $\bar{\gamma}(y) = \gamma^B(y)$, and x, y, and $\left(\hat{\gamma}^F, \hat{\gamma}^W\right)$ contradict the strict stability of (μ^A, γ^A). Thus, $(\bar{\mu}, \bar{\gamma})$ is strictly stable.

Symmetric arguments establish the existence and strict stability of $(\underline{\mu}, \underline{\gamma})$. □

References

[1] Ashlagi, I., Kanoria, Y., and Leshno, J. (Forthcoming.) "Unbalanced Random Matching Markets: the Stark Effect of Competition," *Journal of Political Economy.*

[2] Anonymous (2011). "Too Good to Fire?" Webpage, https://hrguy4omni.wordpress.com/2011/08/22/too-good-to-fire/.

[3] Becker, G. (1973). "A Theory of Marriage: Part I," *Journal of Political Economy*, **81**, pp. 813-846.

[4] Board, S., Meyer-ter-Vehn, M., and Sadzik, T. (2016). "Recruiting Talent," *working paper.*

[5] Campbell III, C. (1993). "Do Firms Pay Efficiency Wages? Evidence with Data at the Firm Level," *Journal of Labor Economics*, **11**, pp. 442-470.

[6] Carmichael, L. and MacLeod, B. (1997). "Gift Giving and the Evolution of Cooperation," *International Economic Review*, **38**, pp. 485–509.

[7] Citanna, A., and Chakraborty, A. (2005). "Occupational choice, incentives and wealth distribution," *Journal of Economic Theory*, **122**, pp. 206-224.

[8] Crawford, V. (1991). "Comparative statics in matching markets," *Journal of Economic Theory*, **54**, pp. 389-400.

[9] Datta, S. (1996). "Building Trust," *working paper*.

[10] Demange, G. and Gale, D. (1985). "The Strategy Structure of Two-Sided Matching Markets," *Econometrica*, **53**, pp. 873-888.

[11] Fujiwara-Greve, T., and Okuno-Fujiwara, M. (2009). "Voluntarily Separable Repeated Prisoner's Dilemma," *Review of Economic Studies*, **76**, pp. 993-1021.

[12] Ghosh, P. and Ray, D. (1996). "Cooperation in Community Interaction without Information Flows," *Review of Economic Studies*, **63**, pp. 491–519.

[13] Gretsky, N., Ostroy, J., and Zame, W. (1992). "The Nonatomic Assignment Model," *Economic Theory*, **2**, pp. 103–127.

[14] Kaya, A., and Vereshchagina, G. (2015). "Moral hazard and sorting in a market for partnerships," *Economic Theory*, **60**, pp. 73-122.

[15] Kaneko, M. (1982). "The Central Assignment Game and the Assignment Markets," *Journal of Mathematical Economics*, **10**, pp. 205–232.

[16] Knuth, D. (1976) *Marriages Stables*. Montreal: Montreal University Press, 1976.

[17] Kranton, R. (1996a). "The Formation of Cooperative Relationships," *Journal of Law, Economics, & Organization*, **12**, pp. 214–233.

[18] Kranton, R. (1996b). "Reciprocal Exchange: A Self-Sustaining System," *American Economic Review*, **86**, pp. 830–851.

[19] Krueger, A., and Summers, L. (1988). "Efficiency Wages and the Inter-Industry Wage Structure," *Econometrica*, **56**, pp. 259-293.

[20] Legros, P., and Newman, A. (2007). "Beauty is a Beast, Frog is a Prince: Assortative Matching with Nontransferabilities," *Econometrica*, **75**, pp. 1073-1102.

[21] McAdams, D. (2011). "Performance and Turnover in a Stochastic Partnership," *American Economic Journal: Microeconomics*, **3**, pp. 107-142.

[22] Morris, R. (2011). "What To Do with the Successful But Lazy Salesperson?" Webpage, http://fistfuloftalent.com/2011/10/what-to-do-with-the-successful-but-lazy-salesperson.html.

[23] Roth, A., Rothblum, U., and Vande Vate, J. (1993). "Stable Matchings, Optimal Assignments, and Linear Programming," *Mathematics of Operations Research,* **18**, pp. 803–828.

[24] Serfes, K. (2005). "Risk sharing and incentives: Contract design under two-sided heterogeneity," *Economics Letters*, **88**, pp. 343-349.

[25] Serfes, K. (2008). "Endogenous matching in a market with heterogeneous principals and agents," *International Journal of Game Theory*, **36**, pp. 587-619.

[26] Shimer, R. and Smith, L. (2000). "Assortative Matching and Search," *Econometrica*, **68**, pp. 343-369.

[27] Watson, J. (1999). "Starting Small and Renegotiation," *Journal of Economic Theory*, **85**, pp. 52-90.

[28] Wright, D. (2004). "The Risk and Incentives Trade-off in the Presence of Heterogeneous Managers," *Journal of Economics*, **88**, pp. 209-223.