I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13328

Taxation in Matching Markets

Arnaud Dupuy
Alfred Galichon
Sonia Jaffe
Scott Duke Kominers

JUNE 2020

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13328

Taxation in Matching Markets

Arnaud Dupuy
CREA, University of Luxembourg and IZA

Alfred Galichon
New York University, Toulouse School of Economics and IZA

Sonia Jaffe
University of Chicago

Scott Duke Kominers
Harvard Business School, Harvard University and NBER

JUNE 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Taxation in Matching Markets*

We analyze the effects of taxation in two-sided matching markets where agents have heterogeneous preferences over potential partners. Our model provides a continuous link between models of matching with and without transfers. Taxes generate inefficiency on the allocative margin, by changing who matches with whom. This allocative inefficiency can be non-monotonic, but is weakly increasing in the tax rate under linear taxation if each worker has negative non-pecuniary utility of working. We adapt existing econometric methods for markets without taxes to our setting, and estimate preferences in the college-coach football market. We show through simulations that standard methods inaccurately measure deadweight loss.

JEL Classification: C78, D3, H2, J3

Keywords: matching, taxation

Corresponding author:
Arnaud Dupuy
CREA
University of Luxembourg
6, rue Richard Coudenhove-Calergi
L-1359 Luxembourg
E-mail: a rnaud.dupuy@uni.lu

* Dupuy gratefully acknowledges the support of a FNR grant C14/SC/8337045. Galichon gratefully acknowledges the support of NSF grant DMS-1716489, and ERC grants FP7-295298, FP7-312503, and FP7-337665. Jaffe gratefully acknowledges the support of an NSF Graduate Research Fellowship and a Terence M. Considine Fellowship in Law and Economics. Kominers gratefully acknowledges the support of NSF grants CCF-1216095 and SES-1459912, an NSF Graduate Research Fellowship, the Ng Fund and the Mathematics in Economics Fund of the Harvard Center of Mathematical Sciences and Applications, the Harvard Milton Fund, a Terence M. Considine Fellowship in Law and Economics, a Yahoo! Key Scientific Challenges Fellowship, and an AMS{Simons Travel Grant.

1 Introduction

Taxation on transfers induces distortion along several margins resulting in a deadweight loss. The bulk of the existing literature has principally studied the extensive margin, i.e. the drop in participation following a tax increase (see e.g. Meyer (2002) and Saez (2002)) and the intensive margin, i.e. the drop in hours of work following a tax increase (Blundell et al. (1998) and Saez (2004)).[1] Often, the impact of these distortions can be fully captured by the elasticity of taxable income (see e.g., Feldstein (1999) and Chetty (2009)) and therefore can be quantified from only wage data when the tax system changes.

In this paper, we study an alternative type of distortion, namely the *matching distortion*, that has been absent in the literature.[2] The matching distortion is the effect of taxation on which workers work for which firms. In the presence of horizontal preference and productivity heterogeneity – workers (firms) disagree about the desirability of different firms (workers) – raising taxes can affect the sorting of workers across firms, thereby affecting the efficiency of the market. Taxes affect the worker-firm sorting because they reduce large transfers more than small ones, so high taxes diminish the extent to which productivity differences are reflected in post-tax wages. Thus, under high taxes, workers may choose the firms they like working for instead of the ones at which they are most productive. An inefficient allocation of workers to firms may not be apparent in the set of workers employed (or the hours that they work). Moreover, while the reallocation affects wages, the efficiency loss from a reallocation is not captured by the changes in wages.

To analyze the matching distortion, we use a framework where agents have idiosyncratic values of matching with potential match partners, as in the matching-without-transfers literature (e.g., Gale and Shapley (1962); Roth (1982)); agents can make transfer payments to their match partners (in the spirit of, e.g. Koopmans and Beckmann (1957); Shapley and Shubik (1971); Becker (1974)). Unlike most matching models, we allow for transfers to be "taxed," causing some of each payment to be taken from the agents.[3] With a proportional

[1] The current literature distinguishes a third type of distortion: the impact of taxation on job search behavior, see e.g. Gentry and Hubbard (2004), Holzner and Launov (2012) and Epstein and Nunn (2013)).

[2] Prior work on taxation and heterogeneous agents, such as occupational choice (e.g., Parker (2003); Sheshinski (2003); Powell and Shan (2012); Lockwood et al. (2017)), Roy models (e.g., Rothschild and Scheuer (2012); Boadway et al. (1991)), and others (e.g., Scheuer and Werning (2017)) only reflects part of the matching distortion we introduce because it does not model either i) the two-sidedness of the market or ii) the fact that, when preferences or productivities are heterogeneous, firms may keep some of the productivity surplus – if a worker is more productive at one firm than at any other, that one firm need not pay the worker his full productivity in equilibrium, in the latter. Explicitly modeling firms allows for the possibility of taxation affecting firms' surplus.

[3] We do not explicitly model the central authority that collects the tax. Our welfare analysis focuses on total match value, implicitly assuming that the social value of tax revenue equals the private value.

tax τ, an agent receives fraction $(1-\tau)$ of the amount his partner gives up (see Section 2).[4] The intermediate tax levels we consider introduce a continuum of models between the well-studied extremes of matching with and without transfers. While prior work has analyzed frameworks that can embed our intermediate transfer models (e.g. Crawford and Knoer, 1981; Kelso and Crawford, 1982; Quinzii, 1984; Hatfield and Milgrom, 2005), it has focused on the structure of the sets of stable outcomes *within* (fixed) models and has not looked at how the efficiency of stable outcomes changes *across* transfer models and is therefore unable to analyze the effect of taxation.[5]

Although our analysis is presented in the language of labor markets, it may also be relevant for other matching markets. Some transfers may be non-monetary and therefore might not be valued equally by givers and receivers: colleges may offer free housing to scholarship students, which might cost them more to provide than students' would be willing to pay for it. Marriage markets also often have in-kind transfers; an agent may value receiving a gift less than it costs his or her partner (in time and money) to give it.[6] Taxation can be reinterpreted as representing the frictions or loss associated to in-kind transfers. In college admissions and marriage markets, positive transfers may flow in both directions.

We show that the classical economic intuition that raising taxes always increases equilibrium deadweight loss holds *if* agents on one side of the market do not match unless they receive positive wages; however, raising taxes can decrease the deadweight loss in markets where transfers flow in both directions. Most labor markets have wages flowing from firms to workers, although there may be internships that workers would pay to get. There are other, more balanced, matching markets, such as the college and marriage examples discussed above, where it may be more reasonable to think of transfers flowing in both directions. In these more symmetric markets, raising taxes can decrease deadweight loss if it prevents an agent from making a large enough (post-tax) transfer to an inefficient match partner, causing the agent to instead match with the (efficient) match partner he would match with (and receive a transfer from) absent taxation (see Example 1). This implies that it is non-trivial to predict the sign of efficiency consequences of a reduction in transfer frictions.

After laying out the general model and theoretical results, we turn to a more specific model for the purposes of estimation. Assuming that a firm's value for workers is separable across matches, the model becomes equivalent to a one-to-one matching model, and we show

[4] In Appendix D we look at lump sum transfers, in which a fixed amount ℓ is subtracted from transfers.

[5] Galichon et al. (2017) show conditions for uniqueness of equilibria under imperfectly transferable utility, along with algorithms for finding equilibria, but do not compare outcomes across different levels of transferability. Legros and Newman (2007) do examine how outcomes change across transfer models, but they use one-dimensional agent types, thus preventing matching distortions.

[6] A similar idea is modeled by Arcidiacono et al. (2011), who treat sexual activity as an imperfect transfer from women to men in the context of adolescent relationships.

how the Choo and Siow (2006) framework for matching with perfect transfers and logit (Gumbel-distributed) heterogeneity in preferences[7] can be adapted to allow for taxation. Categorizing agents by observable type and putting structure on the form of unobserved heterogeneity allows for the identification of match values from data on matching patterns and wages. We also use results from Dupuy and Galichon (2017) for maximum likelihood estimation of amenities and productivities in matching markets when transfers are observed with noise.

As a proof of concept, we apply this technique to estimate job amenities and productivities of college football coaches in the 2013 National Collegiate Athletic Association (Division I). This choice is motivated by several observations. First, coaches' preferences for and productivities at the various colleges are arguably highly differentiated. Second, from an empirical point of view, data about coaches' and colleges' characteristics, coaches' compensation, and teams' performances are readily available. Third, the market for college coaches is large, with football activities generating \$3B in revenues for colleges in this division. Fourth, average yearly salaries of head coaches in Division I are around \$1.8M—well above the cutoff for the top marginal tax bracket in all states, justifying a linear approximation of taxation of coaches' salaries. We use the estimated match values of coaches and teams to simulate the market equilibrium under alternative tax policies. We show that the true change in market surplus is not well approximated by formulas that do not account for the matching nature of the market.

The remainder of the paper is organized as follows: Section 2 introduces our model. Sections 3 presents the theoretical results. Section 4 describes the corresponding econometric model and results on identification. Section 5 describes the data and methods used in our application; it presents the parameter estimates and simulations of counter-factual tax policies. Section 6 concludes. All proofs as well as auxiliary results are presented in the Appendix.

2 General Model

We study a two-sided, many-to-one matching market with fully heterogeneous preferences. We refer to agents on one side of the market as *firms*, denoted $f \in F$; we refer to agents on the other side as *workers*, denoted $w \in W$. Each agent $i \in F \cup W$ derives value from being matched to agents on the other side of the market. We denote these *match values* by $\gamma_{f,D}$ for the value $f \in F$ obtains from matching with the set of workers $D \subseteq W$ and $\alpha_{f,w}$ for the value $w \in W$ obtains from matching with firm $f \in F$. Unmatched agents can

[7]This framework has been extended to general heterogeneity by Galichon and Salanié (2014).

be thought of as being matched to themselves; without loss of generality, we normalize the value of being unmatched (an agent's reservation value) to 0, setting $\gamma_{f,f} = \alpha_{w,w} = 0$ for all $f \in F$ and $w \in W$. In the labor market context, $\gamma_{f,D}$ may represent the productivity of the set of workers D when employed by firm f and $\alpha_{f,w}$ may be the utility or disutility worker w gets from working for f.[8]

Note that it is possible for workers to disagree about the relative desirabilities of potential firms and for firms to disagree about the relative values of potential workers. For our initial results, we impose no structure on workers' match values and only impose enough structure on firms' preferences to ensure the existence of equilibria. For example, the match values could be random draws or may result from an underlying utility or production function in which agents have multi-dimensional types and preferences. To ensure existence of equilibria, we assume that firms' preferences satisfy the standard Kelso and Crawford (1982)/Hatfield and Milgrom (2005) *substitutability* condition: the availability of new workers cannot make a firm want to hire a worker it would otherwise reject.[9]

A *matching* μ is an assignment of agents such that each firm is either matched to itself (*unmatched*) or matched to a set of workers who are matched to it. Denoting the power set of W by $\wp(W)$, a matching is then a mapping μ such that

$$\mu(f) \in (\wp(W) \setminus \emptyset \cup \{f\}) \quad \forall f \in F,$$
$$\mu(w) \in (F \cup \{w\}) \quad \forall w \in W,$$

with $w \in \mu(f)$ if and only if $\mu(w) = f$.

We allow for the possibility of (at least partial) transfers between matched agents. We denote the transfer from f to w by $t_{f,w} \in \mathbb{R}$; if f receives a positive transfer from w, then $t_{f,w} < 0$. A transfer vector t identifies (prospective) transfers between all firm–worker pairs, not just between those pairs that are matched. We also include in the vector t "transfers" $t_{i,i}$ for all agents $i \in F \cup W$, with the understanding that $t_{i,i} = 0$. For notational convenience, we denote by $t_{f,D}$ the total transfer from firm f to workers in $D \subseteq W$:

$$t_{f,D} \equiv \sum_{w \in D} t_{f,w}.$$

In the presence of taxation, an agent might not receive an amount equal to that which

[8]Although it may seem that $\gamma_{f,D}$ should be positive and $\alpha_{f,w}$ should be negative, for our general analysis we do not make sign assumptions. That is, we allow for the possibility of highly demanded internships and for counterproductive employees.

[9]Substitutability plays no role in our analysis other than ensuring, through appeal to previous work (Kelso and Crawford (1982)), that equilibria exist. Thus, we leave the formal discussion of the substitutability condition to Appendix B.

his match partner gives up; the *transfer function* $\xi(\cdot)$ converts an agent's transfer payment into the amount that agent's partner receives. That is, when a firm pays a worker t, that worker receives $\xi(t)$; conversely, when worker pays t to a firm, the firm's post-tax transfer is $\xi^{-1}(t)$.

In the specific case of proportional taxation, the transfer function is

$$\xi_\tau(t_{f,w}) \equiv \begin{cases} (1-\tau)t_{f,w} & t_{f,w} \geq 0 \\ \frac{1}{(1-\tau)}t_{f,w} & t_{f,w} < 0. \end{cases}$$

Figure 1 illustrates the transfer function $\xi_\tau(\cdot)$ for different tax rates τ.

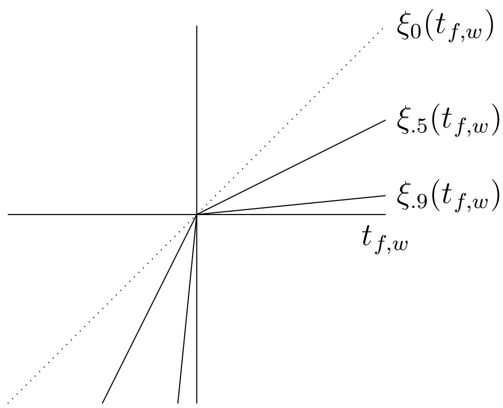


Figure 1: Transfer function $\xi_\tau(\cdot)$.

An *arrangement* $[\mu; t]$ consists of a matching and a transfer vector.[10] We assume that agent payoffs are quasi-linear in transfers and that agents only care about their own match partner(s). With these assumptions, the *payoffs* of arrangement $[\mu; t]$ for firm $f \in F$ and worker $w \in W$ are

$$u_f([\mu;t]) \equiv \gamma_{f,\mu(f)} - t_{f,\mu(f)},$$
$$u_w([\mu;t]) \equiv \alpha_{\mu(w),w} + \xi(t_{\mu(w),w}).$$

The payoff of worker $w \in W$ depends on the transfer function $\xi(\cdot)$. Note that both the match values and the transfers may be either positive or negative. As noted above, without loss of generality, we normalize the payoff of all unmatched agents to 0.

Our analysis focuses on the arrangements that are stable, in the sense that no agent wants to deviate. Because we work with arrangements, which specify full transfer vectors

[10]Here were use the term "arrangement" instead of "outcome" for consistency with the matching literature (e.g., Hatfield et al. (2013)), which uses the latter term when the transfer vector only includes transfers between agents who are matched to each other.

(rather than just transfers associated with the partnerships that arise in equilibrium), we use a stability concept that corresponds to competitive equilibrium; as Kelso and Crawford (1982) showed, this concept is equivalent to the other standard stability concept of matching theory, which rules out the possibility of "blocks" in which groups of agents jointly deviate, while potentially adjusting transfers (see Appendix B).[11]

Definition 1. *An arrangement* $[\mu; t]$ *is* stable given transfer function $\xi(\cdot)$ *if the following conditions hold:*

1. *Each agent (weakly) prefers his assigned match partner(s) (with the corresponding transfer(s)) to being unmatched, that is,*

$$u_i([\mu; t]) \geq 0 \qquad\qquad \forall i \in F \cup W.$$

2. *Each firm (weakly) prefers its assigned match partners (with the corresponding transfers) to any alternative set of workers (with the corresponding transfers), that is,*

$$u_f([\mu; t]) = \gamma_{f,\mu(f)} - t_{f,\mu(f)} \geq \gamma_{f,D} - t_{f,D}, \qquad \forall f \in F \text{ and } D \subseteq W;$$

and each worker (weakly) prefers his assigned match partner (with the corresponding transfer) to any alternative firm (with the corresponding transfer), that is,

$$u_w([\mu; t]) = \alpha_{\mu(w),w} + \xi(t_{\mu(w),w}) \geq \alpha_{f,w} + \xi(t_{f,w}) \qquad \forall w \in W \text{ and } f \in F.$$

We say that a matching μ is *stable given transfer function* $\xi(\cdot)$ if there is some transfer vector t such that the arrangement $[\mu; t]$ is stable given $\xi(\cdot)$; the transfer vector, t, is said to *support* μ (given $\xi(\cdot)$).

The assumption of substitutable preferences ensures that at least one stable arrangement always exists (by results of Kelso and Crawford (1982); see Appendix B for details).

In analyzing stable arrangements we focus on the total value, inclusive of tax revenue,

$$\begin{aligned}
\mathfrak{M}(\mu, t) &\equiv \sum_{i \in F \cup W} u_i(\mu(i), t) + \sum_{f \in F} \left(t_{f,\mu(f)} - \xi(t_{f,\mu(f)})\right) \\
&= \sum_{f \in F} \left(\gamma_{f,\mu(f)} - t_{f,\mu(f)}\right) + \sum_{w \in W} \left(\alpha_{\mu(w),w} + \xi(t_{f,\mu(f)})\right) + \sum_{f \in F} \left(t_{f,\mu(f)} - \xi(t_{f,\mu(f)})\right) \\
&= \mathfrak{M}(\mu),
\end{aligned}$$

[11] This is consistent with the approach taken in most of the empirical matching literature (see, e.g., Chiappori and Salanié (2016)), which uses "stability," "competitive equilibrium," and (often) just "equilibrium" interchangeably. We adopt the specific term "stability" instead of "competitive equilibrium" to simplify the exposition, and to highlight the connection of our work to the broader two-sided matching literature.

which is just the *total match value* and depends only on the matching μ, not on the supporting transfer vector, t, or the transfer function $\xi(\cdot)$.

Definition 2. *We say that a matching $\hat{\mu}$ is* efficient *if it maximizes total match value among all possible matchings, i.e. if $\mathfrak{M}(\hat{\mu}) \geq \mathfrak{M}(\mu)$ for all matchings μ.*[12]

Some of our analysis focuses on markets in which workers have nonpositive valuations for matching, so that they will only match if paid positive "wage" transfers. Formally, we say that a market is a *wage market* if

$$\alpha_{f,w} \leq 0 \tag{2.1}$$

for all $w \in W$ and $f \in F$. The existence of internships notwithstanding, most labor markets can be reasonably modeled as wage markets.

For simplicity, we set our illustrative examples in *one-to-one matching markets*, in which each firm matches to at most one worker; for such markets, we abuse notation slightly by only specifying match values for firm–worker pairs and writing w in place of the set $\{w\}$ (e.g., $\gamma_{f,\{w\}}$ is denoted $\gamma_{f,w}$).

3 Taxation and Mismatch

In addition to the standard effects of decreasing hours worked or labor market participation, in matching markets, taxes can decrease efficiency by creating mismatch in which workers work for which firms. Somewhat surprisingly, this mismatch is not necessarily monotonic in the tax rate, as illustrated by the examples in the following subsection.

3.1 Some examples

Example 1 (Non-wage market with linear taxation)**.** *We take a simple market with one firm, $F = \{f_1\}$, two workers, $W = \{w_1, w_2\}$, and match values as pictured in Figure 2a. Worker w_1 receives a high value from matching with f_1. Firm f_1 is indifferent towards worker w_1 and receives moderate value from matching with w_2. Worker w_2 has a mild preference for being unmatched, rather than matching with f_1. We can think of w_1 as an intern who*

[12] An alternative welfare measure would be *total agent value*, i.e. total match value minus total tax revenue. However, while government expenditures may not always be valued dollar-for-dollar, including government revenue in welfare is typically considered a better approximation than assigning it no value (Mas-Colell et al., 1995). Moreover, total agent value depends on the transfer vector; as there are frequently many transfer vectors that support a given stable matching, total agent value is not typically well-defined, even fixing a given stable matching and tax function.

would not be very productive in working for f_1, but would learn a lot; w_2 represents a normal worker, who is productive but does not like working. The tax represents an income tax, which the firm also pays on transfers from the intern.

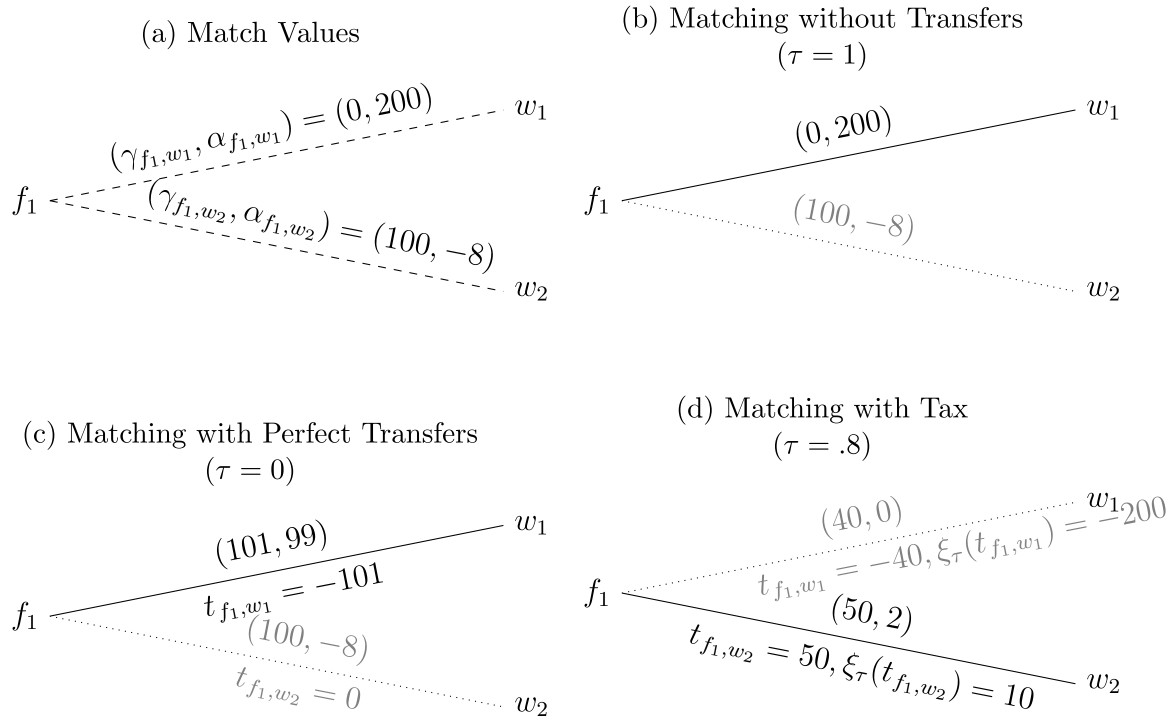


Figure 2: Example 1 – Non-monotonicity under a proportional tax on transfers.

Note: Utilities, net of transfers, are above the lines (firm's, worker's). Possible supporting transfers (when applicable) are below the lines. Solid lines indicate the stable matching.

As illustrated in Figure 2b, when $\tau = 1$ (or when transfers are not allowed), the only stable matching $\hat{\mu}$ has $\hat{\mu}(f_1) = w_1$. Since $\hat{\mu}$ is efficient matching, it is also stable when $\tau = 0$, as shown in Figure 2c. The total match value of $\hat{\mu}$ is $\mathfrak{M}(\hat{\mu}) = 200$. Figure 2d shows that for $\tau = .8$, an inefficient matching $\check{\mu}$, for which $\check{\mu}(f_1) = w_2$, is stable. The inefficient matching generates a total match value $\mathfrak{M}(\check{\mu}) = 92$. Not only is an inefficient matching stable under tax $\tau = .8$, but the efficient matching $\hat{\mu}$ is not stable under this tax, or any tax $\tau \in (.6, .9)$.[13]

While Example 1 may appear quite specialized, simulations suggest that non-monotonicities in the total match value of stable matchings as a function of τ can be relatively common. In

[13]For that range, $(100 - 200(1-\tau))(1-\tau) - 8 > 0$, so that the maximum f_1 can transfer to w_2 while still preferring w_2 to w_1 is sufficient to outweigh the disutility w_2 gets from matching to f_1. Note that here total agent payoffs (match value minus government revenue), like total match value, can be non-monotonic. When $\tau = 1$, total agent value is 200 (assuming they do not burn money); when $\tau = .8$ it is 52.

simulations of small markets with utilities drawn independently from a uniform distribution on $[-.5, .5]$, we find that 55% of markets exhibit non-monotonicities. (See Appendix C for details.) On average, the value drop at a non-monotonicity is 12% of the difference between the optimal match and the worst match that is stable at any tax rate (for that market). While our simulations suggest non-monotonicities in the tax rate are not just artifacts of the example selected, they also suggest that non-monotonicities are relatively rare at more realistic tax rates ($\tau \in [0, .5)$) and tend not to persist over large ranges of τ.

Although the total match value of stable matchings may decrease when the tax rate falls, an arrangement that is stable under a tax rate $\hat{\tau}$ must improve the payoff of at least one agent, relative to an arrangement that is stable under a tax rate $\check{\tau} > \hat{\tau}$ – raising the tax rate cannot lead to a Pareto improvement.[14] These non-monotonicities can only arise in non-wage markets, so are perhaps more relevant to student-college matching markets or marriage markets – where transfer frictions can act like taxes – than to labor markets.

Example 2 shows that, if taxes are nonlinear, lower taxes can lower the match value even in wage markets.

Example 2 (Wage market with piecewise linear taxation)**.** *Figure 3a shows the match values for a market with two firms, $F = \{f_1, f_2\}$, two workers, $W = \{w_1, w_2\}$. Worker w_1 (w_2) is fairly productive and receives moderate disutility working for firm f_1 (f_2). Firm f_1 could also hire worker w_2 who is much more productive, but dislikes working for f_1 more than worker w_1 does. There is no surplus from f_2 hiring w_1.*

Figure 3b shows the transfer functions. There is a marginal tax rate $\tau_1 = .5$ on transfers up to 20 and we consider marginal tax rates of $\tau_2 \in \{.5, .75\}$ on the part of the transfer above 20. Fig 3c shows the equilibrium with the lower marginal tax rate of $\tau_2 = .5$. The only stable match is $\mu(f_1) = w_2$, which gives a match utility of 20 and is inefficient. Fig 3d shows the equilibrium when $\tau_2 = .75$. The stable match is $\mu(f_1) = w_1$ and $\mu(f_2) = f_2$, which gives total a total match value of 22 and is efficient. Raising the marginal tax rate from $\tau_2 = .5$ to $\tau_2 = .75$ raises *the match utility.*

If tax rates differ across agents – e.g., states have different income taxes – then we arrive at a similar situation as under the non-linear taxes in Example 2. Lowering any given tax rate could disproportionately encourage a less efficient matching, decreasing welfare. Even if a tax rate is lowered to make the tax rates across agents more equal, that can still lower welfare.

[14] See Appendix A and B for the intuition and proof.

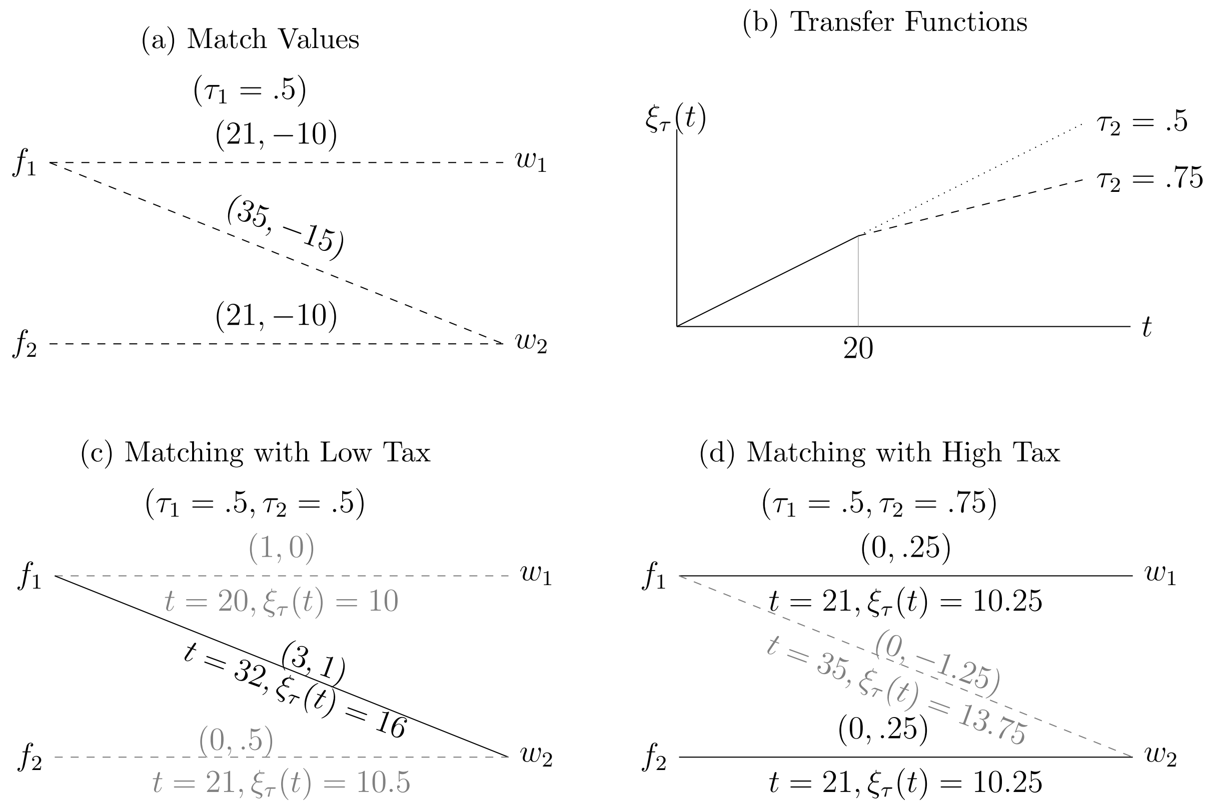


Figure 3: Example 2 – Non-monotonicity under a lump sum tax on transfers.

Note: Utilities, net of transfers, are above the lines (firm's, worker's). Possible supporting transfers (when applicable) are below the lines. Solid lines indicate the stable matching.

3.2 Monotonicity

Despite our negative results, we show that wage markets and proportional taxation together are sufficient to ensure that total match value is (weakly) monotonic in the tax rate. Since payments in wage markets flow from firms to workers, any stable matching can be supported by a non-negative transfer vector.[15] Thus, the transfer function with proportional taxation $\xi_\tau(\cdot)$ takes the simpler form

$$\xi_\tau(t_{f,w}) = (1 - \tau)t_{f,w} \geq 0.$$

Since all positive transfers are paid from firms to workers, there cannot be a scenario in which, as in Example 1, when the tax is reduced, a firm can transfer enough to get a worker

[15]There may be a supporting transfer vector where some off-path transfers (transfers between unmatched agents) are negative, but in that case there is always another supporting transfer vector that replaces those negative transfers with 0s. Our results only require the existence of a non-negative supporting transfer vector.

it prefers (w_2), but when the tax falls more, a different worker (w_1) can "buy back" the firm.[16] Moreover, as taxes are linear, all transfers are equally (proportionally) affected by the tax rate. It cannot be the case, as in Example 2, that, some transfers are more affected by the initial tax, but others are more affected by the tax increase.

Theorem 1. *In a wage market with proportional taxation, a decrease in taxation (weakly) increases the total match value of stable matchings. That is, if in a wage market matching* $\check{\mu}$ *is stable under tax* $\check{\tau}$*, matching* $\hat{\mu}$ *is stable under tax* $\hat{\tau}$*, and* $\hat{\tau} < \check{\tau}$*, then*

$$\mathfrak{M}(\hat{\mu}) \geq \mathfrak{M}(\check{\mu}).$$

To prove Theorem 1, we let $\hat{t} \geq 0$ and $\check{t} \geq 0$ be transfer vectors supporting $\hat{\mu}$ and $\check{\mu}$ respectively. The stability of $[\hat{\mu}; \hat{t}]$ under tax $\hat{\tau}$ implies that

$$\gamma_{f,\hat{\mu}(f)} - \hat{t}_{f,\hat{\mu}(f)} \geq \gamma_{f,\check{\mu}(f)} - \hat{t}_{f,\check{\mu}(f)}, \text{ and} \tag{3.1}$$

$$\alpha_{\hat{\mu}(w),w} + (1-\hat{\tau})\hat{t}_{\hat{\mu}(w),w} \geq \alpha_{\check{\mu}(w),w} + (1-\hat{\tau})\hat{t}_{\check{\mu}(w),w}. \tag{3.2}$$

Summing (3.1) and (3.2) across agents, using the fact that the total transfers paid by all firms equals the total transfers paid by all workers' match partners, that is

$$\sum_{f \in F} t_{f,\mu(f)} = \sum_{f \in F} \sum_{w \in \mu(f)} t_{f,w} = \sum_{w \in W} t_{\mu(w),w}, \tag{3.3}$$

and regrouping terms, we find that

$$\begin{aligned} \mathfrak{M}(\hat{\mu}) - \mathfrak{M}(\check{\mu}) &= \sum_{f \in F} (\gamma_{f,\hat{\mu}(f)} - \gamma_{f,\check{\mu}(f)}) + \sum_{w \in W} (\alpha_{\hat{\mu}(w),w} - \alpha_{\check{\mu}(w),w}) \\ &\geq \hat{\tau} \sum_{f \in F} \left(\hat{t}_{f,\hat{\mu}(f)} - \hat{t}_{f,\check{\mu}(f)} \right). \end{aligned} \tag{3.4}$$

Equation (3.4) shows that $\hat{\mu}$ has higher match utility than $\check{\mu}$ if, on average, the transfers to an agent's match partners under $\hat{\mu}$ is greater than the off-path transfer to his partner under $\check{\mu}$. Intuitively, the match-partner transfers must be larger than the off-paths transfers since

[16] The non-monotonicity in Example 1 arises from transfers flowing in both directions, either simultaneously or across equilibria. As transfers are an equilibrium phenomenon, requiring that transfers flow in one direction does not directly correspond to conditions on the primitives of the market. However, the wage market condition we use in Theorem 1 is a *sufficient* condition on primitives to guarantee that transfers flow in one direction, and thus is sufficient to rule out non-monotonicity. All the results in this section hold in any market where transfers always (across stable allocations and tax rates) flow in one direction, even if it is not a wage market.

the tax change has a larger effect on larger transfers. Assume we had

$$\sum_{f \in F} \left(\hat{t}_{f,\hat{\mu}(f)} - \hat{t}_{f,\check{\mu}(f)} \right) < 0.$$

Then lowering the tax from $\check{\tau}$ to $\hat{\tau}$ would increase workers' relative preference for $\check{\mu}$ over $\hat{\mu}$. Since $\hat{\mu}$ is stable under the lower tax $\hat{\tau}$, the difference in (3.4) must be positive, implying Theorem 1.

Although total match value in wage markets increases as the tax is reduced, *individual* payoffs may be non-monotonic. For example, pursuant to a tax decrease, a firm f may be made worse off because his match partner is now able to receive more from some other firm. Firm f might lose his match partner to his competitor; even if f's match is unchanged, its total payoff may decrease because it is forced to increase its transfer to compensate for a competitor's increased offer.

3.3 Renormalizing Utilities in Wage Markets

In a wage market, we can renormalize worker utilities in order to express them in *pre-tax dollars*, by defining

$$u_w^\tau([\mu, t]) \equiv \frac{1}{1-\tau} u_w([\mu, t]) = \frac{1}{1-\tau} \alpha_{\mu(w),w} + t_{\mu(w),w}.$$

Since the firms care about pre-tax dollars, putting workers' match values in pre-tax dollars makes them easier to aggregate with firms' match values. A post-tax dollar is equivalent to $\frac{1}{1-\tau}$ pre-tax dollars, so workers' match utilities must be divided by $(1-\tau)$ to 'pre-tax' values. Workers' *relative* preferences, and therefore the outcomes in the market, are unchanged by the renormalization. Since the outcomes are unchanged by the renormalization, the matching market with tax rate τ has the same set of stable matchings as a market without taxes that has match values $\{\tilde{\alpha}_{f,w}\} = \frac{1}{1-\tau}\{\alpha_{f,w}\}$ and $\{\gamma_{f,w}\}$. We can then use results from the literature on matching with transfers to characterize the stable matchings.

We know that with (perfect) transfers, only efficient matchings – that maximize the sum of agents' payoffs – are stable; this result no longer holds with a nonzero tax rate because we are trying to add apples (firm's utilities $\{\gamma_{f,w}\}$, which are denominated in pre-tax dollars) to oranges (worker's utilities $\{\alpha_{f,w}\}$, which are denominated in after-tax dollars). If we express everything in pre-tax dollars, however, we get a result parallel to the usual efficiency result: the equilibrium matching maximizes the sum of output measured in a common denomination.

Proposition 1. *In a wage market with proportional taxation, if a matching $\check{\mu}$ is stable under tax $\check{\tau}$, then it is a matching that maximizes the sum of firm utilities plus $\frac{1}{1-\check{\tau}}$ times the sum* *of worker utilities,*[17]

$$\check{\mu} \in \arg\max_{\{\mu\}} \left[\sum_{f \in F} \gamma_{f,\mu(f)} + \frac{1}{1-\check{\tau}} \sum_{w \in W} \alpha_{\mu(w),w} \right].$$

Additionally:

1. *Workers' match value,* $\sum_{w \in W} \alpha_{\mu(w),w}$*, is weakly increasing in the tax rate.*

2. *Firms' match value,* $\sum_{f \in F} \gamma_{f,\mu(f)}$*, is weakly decreasing in the tax rate.*

Absent taxation, stable matchings maximize the sum of match values with equal weight on firms and workers; under taxation, stable matchings still maximize the sum of match values, but with different weights. Taxes decrease the relative weight put on firms' preferences because their ability to express those preferences to workers via wages is decreased. As a result, the overall match value of firms decreases with the tax rate. Conversely, as their preferences get relatively more weight, the overall match value of workers increases with the tax rate. However, workers are still better off under lower taxes if they receive sufficiently higher transfers so as to more than compensate for their lower match values.

In Appendix A we show some additional features of wage markets:

- There is some $\underline{\tau}$ such that only an efficient matching is stable for $\tau < \underline{\tau}$.

- If two distinct matchings $\check{\mu}$ and $\hat{\mu}$ are both stable under tax τ, then they can be supported by the same transfer vectors and, for any supporting transfer vector, all agents are indifferent between the two allocations (based on Hatfield et al., 2013).[18]

- If for any τ there are multiple stable arrangements, then firms and workers have opposing preferences. If all firms prefer $[\hat{\mu}; \hat{t}]$ to $[\hat{\mu}; t]$, then all workers prefer $[\hat{\mu}; t]$ to $[\hat{\mu}; \hat{t}]$.

[17]Note that this result still holds if the tax rates are individual specific. In that case one has

$$\check{\mu} \in \arg\max_{\{\mu\}} \left[\sum_{f \in F} \gamma_{f,\mu(f)} + \sum_{w \in W} \frac{1}{1-\check{\tau}_w} \alpha_{\mu(w),w} \right].$$

[18]The difference in total match value of the two matchings equals the difference in revenue for the two stable matchings, for a given supporting transfer vector. Unfortunately this equality does not give much traction empirically because as the tax rate changes, transfers will change even when the underlying match does not change (so there is no change in total match value). Also, even at the tax rate where multiple matchings are stable, there may be multiple supporting transfer vectors and the revenue between $[\hat{\mu}, \hat{t}]$ and $[\check{\mu}, \check{t}]$ does not tell us anything about the difference in total match value between $\hat{\mu}$ and $\check{\mu}$.

- If there is more than one tax rate under which two distinct matchings $\check{\mu}$ and $\hat{\mu}$ are both stable, then $\check{\mu}$ and $\hat{\mu}$ must have the same total match value, $\mathfrak{M}(\check{\mu}) = \mathfrak{M}(\hat{\mu})$. Moreover, firms are indifferent in aggregate between the matchings. If $\check{\mu}$ and $\hat{\mu}$ do not have the same total match value or firms are not indifferent, the only tax rate τ under which they *could* both be stable is

$$\tau = 1 + \frac{\sum_{w \in W} \left(\alpha_{\check{\mu}(w),w} - \alpha_{\hat{\mu}(w),w}\right)}{\sum_{f \in F} \left(\gamma_{f,\check{\mu}(f)} - \gamma_{f,\hat{\mu}(f)}\right)}.$$

  This also implies that for each tax rate the stable matching is generically unique.

In Appendix D, we also discuss lump-sum taxation under which, instead of taking a fixed proportion of any transfer, the tax takes a fixed amount from each transfer. Just as under proportional taxation, with lump-sum taxation there is a possibility for non-monotonicity, and strict wage markets (where worker match values are strictly negative, instead of just non-positive) are needed in order to guarantee monotonicity.

3.4 Implications for Tax Analysis

In addition to causing some workers not to work, taxation generates deadweight loss by changing the matching of workers to firms. The preference heterogeneity (and implied imperfectly elastic supply of jobs) means that firms can have positive surplus; workers' decisions on where to work affect firms' productivity and the opportunities available to other workers. These externalities mean that, unlike in the framework of Feldstein (1999), the deadweight loss cannot be calculated from the change in taxable income.[19]

If workers are paid their productivity, the Feldstein (1999) formula says that the deadweight loss is the product of the tax rate and the change in taxable income:

$$\frac{d\text{DWL}}{d\tau} = \tau \frac{d\text{Taxable Income}}{d\tau}.$$

This formula does not hold in matching markets. There is not a closed-form representation of deadweight loss in matching markets, so we cannot prove that the formula is always wrong, but in Section 5.3 we show that it does not hold in a market with the utilities that we estimate. Intuitively, in matching markets wages can change differently from welfare in two ways. Sometimes when workers switch jobs, their wages drop, but the workers like their jobs correspondingly more. The change in taxable income does not capture the fact that workers like their jobs more, leading the wage-based estimate of DWL to be potentially

[19] Chetty (2009) gives other conditions under which the Feldstein (1999) formula does not hold.

biased upward. However, there can also be increases in wages that reflect lost profits of the firm, rather than increased productivity, so the estimate can also be biased downward.

4 Econometric Framework

The preceding results hold for any formulation of match values. However they do not tell us anything about the magnitude of the distortion from taxation, as the magnitudes depend on the distribution of match values in the market. To estimate match values so we can simulate the effect of taxation, we now adopt the Choo and Siow (2006) structure, which assumes that agents have observable types and limits the role of unobserved heterogeneity in preferences. We extend the Choo and Siow (2006) framework to account for taxes; we allow for tax rates that vary by individual type, to accommodate state taxes and different filing statuses. We assume the market is a wage market and the matching is one-to-one. The latter assumption may appear restrictive, but if firms hire more than one worker and have additively separable preferences, then each firm can be split into as many positions that it has to offer and the model can be recast as a one-to-one model (Roth and Sotomayor (1989)).

We assume that each worker has a (multidimensional) type $x_w \in X$ and each firm f has a (multidimensional) type $y_f \in Y$. For example, a worker's type may include education, age, and ability; a firm's type may include the firm's size or location, its technology, and its management style. There are r_x workers of type x and m_y firms of type y. Match values have a systematic component which depends only on the agents' types, and an additively separable random component that is drawn for each agent-type pair:

$$\begin{aligned}\alpha_{f,w} &= \alpha_{y_f,x_w} + \sigma^W \varepsilon_{y_f,w}\\ \gamma_{f,w} &= \gamma_{y_f,x_w} + \sigma^F \eta_{f,x_w},\end{aligned}$$

where $\varepsilon_{y_f,w}$ and η_{f,x_w} are the heterogeneous, random components of the match utility, drawn from a standard type-I Gumbel distribution[20]; σ^W and σ^F measure the variance in the random components of agents' preferences.

As the number of agents of each type gets large, instead of keeping track of which individuals are matched, a matching can be described by the number of matches between each pair of types, which is unaffected by specific draws of $\varepsilon_{y_f,w}$ and η_{f,x_w}. The equilibrium transfers will also be independent of the random utility draws. The set of feasible matchings, denoted

[20]The Gumbel distribution yields a logit structure. It is possible to generalize the argument to arbitrary distribution of the random components using the methodology of Galichon and Salanié (2014).

$\mathcal{M}(m,r)$, is the set of vectors $\mu \geq 0$ such that

$$\sum_{x\in X} \mu_{y,x} \leq m_y \quad \forall y \in Y$$
$$\sum_{y\in Y} \mu_{y,x} \leq r_x \quad \forall x \in X.$$

Transfers are also unaffected by specific draws of $\varepsilon_{y_f,w}$ and η_{f,x_w} and must be the same for any agents of the same type, $t_{f,w} = t_{y_f,x_w}$.

We normalize the systematic utility of being unmatched to 0 for all worker and firm types and let $\varepsilon_{0,w}$ and $\eta_{f,0}$, also drawn from a standard type-I Gumbel distribution, be the random components for workers and firms, respectively. Agents' utilities are

$$u_w = \max\{\max_y\{\alpha_{y,x_w} + (1-\tau^W_{x_w})(1-\tau^W_{y_f})t_{y,x_w} + \sigma^W\varepsilon_{y,w}\}, \sigma^W\varepsilon_{0,w}\}$$
$$v_f = \max\{\max_x\{\gamma_{y_f,x} - (1+\tau^F_{y_f})t_{y_f,x_w} + \sigma^F\eta_{f,x}\}, \sigma^F\eta_{f,0}\},$$

where $\tau^W_{y_f}$ is an income tax that varies by firm ("state income tax") and $\tau^W_{x_w}$ is an income tax that may vary by worker ("federal income tax") and $\tau^F_{y_f}$ is a payroll tax ("state and federal payroll taxes") that may vary by firm.[21]

4.1 Renormalization

Since we allow for both income and payroll taxes as well as allowing taxes to vary by worker or firm, the renormalization is slightly more complicated than in the case with a single tax rate. Let $\lambda^W_x = \frac{1}{1-\tau^W_x}$ and $\lambda^F_y = \frac{1-\tau^W_y}{1+\tau^F_y}$. Using $x = x_w$ and $y = y_f$, we rescale the amenity and productivity terms, as in Section 3.3, to define

$$\tilde{\alpha}_{y,x} \equiv \lambda^W_x \alpha_{y,x}, \qquad \tilde{\gamma}_{y,x} \equiv \lambda^F_y \gamma_{y,x},$$
$$\tilde{u}_w \equiv \lambda^W_x u_w, \qquad \tilde{v}_f \equiv \lambda^F_y v_y,$$

and

$$\tilde{\phi}_{y,x} \equiv \tilde{\alpha}_{y,x} + \tilde{\gamma}_{y,x} = \lambda^W_x \alpha_{y,x} + \lambda^F_y \gamma_{y,x}.$$

[21]Note that tax rates are now allowed to vary by types of workers and firms to reflect the situation in the data, and while are theoretical results are not guaranteed to hold under those conditions, we do find in our empirical application that welfare is decreasing in tax rates. Note further that for simplicity, we do not include payroll taxes that vary by worker, but the renormalization we present is still possible. Finally, note that under the Tax Cuts and Jobs Act introduced in 2017 in the US, state taxes are no longer deductible from income for federal purposes. As a result, the after-tax income specification $(1-\tau^W_{x_w})(1-\tau^W_{y_f})t_{y,x_w}$ used in the definition of workers' utility no longer applies and should be replaced by $(1-\tau^W_{x_w}-\tau^W_{y_f})t_{y,x_w}$. In the empirical application we use data from 2013, justifying our chosen specification of taxes.

Utilities in the fictitious market, in terms of the rescaled amenity and productivity terms, are given by

$$\tilde{u}_w = \max\{\max_y\{\tilde{\alpha}_{y,x_w} + \tilde{t}_{y,x_w} + \lambda_x^W \sigma^W \varepsilon_{y,w}\}, \lambda_x^W \sigma^W \varepsilon_{0,w}\}$$
$$\tilde{v}_f = \max\{\max_x\{\tilde{\gamma}_{y_f,x} - \tilde{t}_{y_f,x} + \lambda_y^F \sigma^F \eta_{f,x}\}, \lambda_y^F \sigma^F \eta_{f,0}\},$$

where $\tilde{t} = (1-\tau_y^W)t$.

Using the logit distribution of errors, the expected utilities (conditional on observable types) in our constructed market are

$$\tilde{a}_x = \lambda_x^W \sigma^W \log\left(1 + \sum_y \exp\left(\frac{\tilde{\alpha}_{y,x} + \tilde{t}_{y,x}}{\lambda_x^W \sigma^W}\right)\right)$$
$$\tilde{b}_y = \lambda_y^F \sigma^F \log\left(1 + \sum_x \exp\left(\frac{\tilde{\gamma}_{y,x} - \tilde{t}_{y,x}}{\lambda_y^F \sigma^F}\right)\right).$$

Based on the logit conditional choice probabilities, the match frequencies are:

$$\mu_{y,x} = r_x \frac{\exp\left(\frac{\tilde{\alpha}_{y,x} + \tilde{t}_{y,x}}{\lambda_x^W \sigma^W}\right)}{\exp\left(\frac{\tilde{a}_x}{\lambda_x^W \sigma^W}\right)}, \qquad \mu_{y,x} = m_y \frac{\exp\left(\frac{\tilde{\gamma}_{y,x} - \tilde{t}_{y,x}}{\lambda_y^F \sigma^F}\right)}{\exp\left(\frac{\tilde{b}_y}{\lambda_y^F \sigma^F}\right)},$$

from workers' and firms' perspectives, respectively. If we solve for the match frequencies and the transfers in the system, we find that

$$\mu_{y,x} = \left(m_y^{\lambda_y^F \sigma^F} r_x^{\lambda_x^W \sigma^W}\right)^{\frac{1}{\lambda_y^F \sigma^F + \lambda_x^W \sigma^W}} \exp\left(\frac{\tilde{\alpha}_{y,x} + \tilde{\gamma}_{y,x} - \tilde{a}_x - \tilde{b}_y}{\lambda_x^W \sigma^W + \lambda_y^F \sigma^F}\right), \tag{4.1}$$

and

$$\tilde{t}_{y,x} = \frac{\lambda_x^W \sigma^W \left(\tilde{\gamma}_{y,x} - \tilde{b}_y\right) - \lambda_y^F \sigma^F \left(\tilde{\alpha}_{y,x} - \tilde{a}_x\right) + \lambda_x^W \sigma^W \lambda_y^F \sigma^F \log(m_y/r_x)}{\lambda_x^W \sigma^W + \lambda_y^F \sigma^F}. \tag{4.2}$$

Given renormalized match values $\{\tilde{\alpha}_{y,x}, \tilde{\gamma}_{y,x}\}_{y,x}$ and variances σ^F, σ^W, we can solve for the match probabilities and wages. Or given wages and match probabilities, we can estimate the match values and variances.

4.2 Without unmatched agents

In most labor market applications, data on unmatched agents are not readily available. If one only observes matched agents, the same framework applies, using "in-market" match probabilities (the probability of matching with a given partner, conditional on being matched). Instead of the expected utility formulas above, we use

$$
\begin{aligned}
\tilde{a}_x &= \lambda_x^W \sigma^W \log\left(\sum_y \exp\left(\frac{\tilde{\alpha}_{y,x} + \tilde{t}_{y,x}}{\lambda_x^W \sigma^W}\right)\right) \\
\tilde{b}_y &= \lambda_y^F \sigma^F \log\left(\sum_x \exp\left(\frac{\tilde{\gamma}_{y,x} - \tilde{t}_{y,x}}{\lambda_y^F \sigma^F}\right)\right).
\end{aligned}
$$

The formulas for match probabilities and transfers are the same, usings these "in-market" values.

4.3 Maximum Likelihood

We assume that data is a random sample of matched pairs independently drawn from the population of n firm-worker matched pairs indexed by $j \in J$. For each match j, denote by y_j the vector of observed attributes of the firm, x_j the vector of observed attributes of the worker, and t_j the noisy measure of the true salary t_{y_j,x_j}.

Without data on unmatched agents, one cannot estimate the direct effects of own characteristics on own payoffs – the effect of y on $\gamma_{y,x}$ or the effect of x on $\alpha_{y,x}$ – from the matching pattern or wages. So we estimate the effect that each partner's characteristics have on an agent's own match value, both directly and interacted with own characteristics. We parameterize the job amenities and firms' productivity linearly as

$$
\begin{aligned}
\alpha_{y,x} &= \alpha_{y,x}(A) = x^T A_0 y + A_1^T y + A_2^T y^{(2)} \\
\gamma_{y,x} &= \gamma_{y,x}(\Gamma) = x^T \Gamma_0 y + \Gamma_1^T x + \Gamma_2^T x^{(2)},
\end{aligned}
$$

where A_0 and Γ_0 are matrices of coefficients on the cross-terms – the interactions between the worker and firm characteristics, A_1, A_2, Γ_1, and Γ_2 are vectors of coefficients for the direct effects, and $v^{(2)}$ indicates a vector of the same length as v with each term squared.

For a given value of the parameters, the rescaled preferences and productivity are

$$
\begin{aligned}
\tilde{\alpha}_{y,x}(A) &= \lambda_x^W \alpha_{y,x}(A) \\
\tilde{\gamma}_{y,x}(\Gamma) &= \lambda_y^F \gamma_{y,x}(\Gamma) \\
\tilde{\phi}_{y,x}(A, \Gamma) &= \tilde{\alpha}_{y,x}(A) + \tilde{\gamma}_{y,x}(\Gamma).
\end{aligned}
$$

The scaling factors of the unobserved heterogeneity are also rescaled – from σ^W to $\lambda_x^W \sigma^W$ and σ^F to $\lambda_y^F \sigma^F$ respectively. Hence, the distributions of unobserved heterogeneity in the rescaled economy are fundamentally heteroskedastic when λ_x^W and λ_y^F vary with x and y.

Our estimation strategy builds on the Dupuy and Galichon (2017) method for maximum likelihood estimation of amenities and productivities when transfers are observed with noise.[22] Dupuy and Galichon (2017) showed that the log-likelihood function in their context can be written as the sum of two terms, one capturing the likelihood of the observed matchings and one capturing the likelihood of the observed wages.

For a given (A, Γ), we calculate the expected utilities $\tilde{b}_j(A, \Gamma)$ and $\tilde{a}_j(A, \Gamma)$ such that the sum across workers of a firm's probability of matching with that worker must be one for every firm, and similarly for workers: the solution of

$$\begin{cases} \sum_{j' \in J} \exp\left(\frac{\tilde{\phi}_{y_j, x_{j'}}(A,\Gamma) - \tilde{a}_{j'} - \tilde{b}_j}{\lambda^W_{x_{j'}}\sigma^W + \lambda^F_{y_j}\sigma^F}\right) = 1 \quad \forall j \in J \\ \sum_{j' \in J} \exp\left(\frac{\tilde{\phi}_{y_{j'}, x_j}(A,\Gamma) - \tilde{a}_j - \tilde{b}_{j'}}{\lambda^W_{x_j}\sigma^W + \lambda^F_{y_{j'}}\sigma^F}\right) = 1 \quad \forall j \in J \end{cases}$$

with an arbitrarily chosen normalization of $\tilde{a}_{x^*} = 0$. The log-likelihood of the observed match is then

$$\log L_1(A, \Gamma) = \log\left(\prod_{j \in J} \exp\left(\frac{\tilde{\phi}_{y_j, x_j}(A, \Gamma) - \tilde{a}_j - \tilde{b}_j}{\lambda^W_{x_j}\sigma^W + \lambda^F_{y_j}\sigma^F}\right)\right) = \sum_{j \in J}\left(\frac{\tilde{\phi}_{y_j, x_j}(A, \Gamma) - \tilde{a}_j - \tilde{b}_j}{\lambda^W_{x_j}\sigma^W + \lambda^F_{y_j}\sigma^F}\right).$$

For wages, assume that the true (adjusted) wage is observed with error

$$t_j = \tilde{t}_{y_j, x_j}(A, \Gamma, c) + (1 - \tau^W_{y_j})\delta_j,$$

where c is a constant that accounts for the normalization $\tilde{a}_{x^*} = 0$ and δ_j is the measurement error, which follows a $\mathcal{N}(0, s^2)$ distribution, independent of (y_j, x_j). With (4.2), we see that

$$\tilde{t}_{y_j, x_j}(A, \Gamma, c) = \frac{\lambda^W_{x_j}\sigma^W}{\lambda^W_{x_j}\sigma^W + \lambda^F_{y_j}\sigma^F}\left(\tilde{\gamma}_{y_j, x_j} - \tilde{b}_j\right) - \frac{\lambda^F_{y_j}\sigma^F}{\lambda^W_{x_j}\sigma^W + \lambda^F_{y_j}\sigma^F}\left(\tilde{\alpha}_{y_j, x_j} - \tilde{a}_j\right) + c. \tag{4.3}$$

The likelihood of the observed wages is

$$\log L_2\left(A, \Gamma, c, s^2\right) = -\sum_{j=1}^{n}\left(\frac{t_j - \tilde{t}_{y_j, x_j}(A, \Gamma, c)}{(1 - \tau^W_{y_j})}\right)^2 \frac{1}{2s^2} - \frac{n}{2}\log s^2.$$

4.3.1 Productivity

If available, data on firm productivity can also be incorporated, yielding a third term in the expression of the log-likelihood. For our application, we do not have a good measure

[22] Note that, in this setting, each worker (firm) defines his (its) own type.

of teams absolute productivity, but the rankings give us a (noisy) measure of their *relative* productivity: the teams compete with each other and therefore lose more when the other teams are more productive. We add a term to the likelihood that measures the probability of the observed ranking of firms.

In addition to the productivity term γ_{y_j,x_j}, which affects wages and matching, a firm's full productivity is also directly affected by its own characteristics; with another vector of parameters, Γ_D^T, we have $\gamma_j^{\text{Tot}} = \gamma_{y_j,x_j} + \Gamma_D^T y_j$.[23] Let cardinal productivity underlying the rankings be

$$Z_j^* = \gamma_j^{\text{Tot}} + \nu_j,$$

where ν_j are drawn from a type-I Gumbel distribution with scaling parameter $1/\beta$, independently of (y_j, x_j, δ_j). For notational simplicity, we sort matchings by decreasing order of measured performance, so a team's index j is equal to its rank. The probability of the observed ordinal ranking of firms is

$$\begin{aligned}\Pr\left(Z_1^* > Z_2^* > ... > Z_n^* \mid \Gamma, \beta\right) &= \prod_{j=1}^{|J|-1} \Pr\left(\gamma_j^{\text{Tot}} + \nu_j > \max_{j<j'}\left(\gamma_j^{\text{Tot}} + \nu_j\right)\right) \\ &= \prod_{j=1}^{|J|-1} \frac{\exp\left(\beta\gamma_j^{\text{Tot}}\right)}{\sum_{j'\geq j}\exp\left(\beta\gamma_{j'}^{\text{Tot}}\right)}.\end{aligned}$$

The resulting log-likelihood is

$$\log L_3\left(\Gamma, \beta\right) = \sum_{j=1}^{|J|-1}\left(\beta\gamma_j^{\text{Tot}} - \log\sum_{j'\geq j}\exp\left(\beta\gamma_{j'}^{\text{Tot}}\right)\right),$$

where the information from the Zs is reflected in the ordering of the js.

Finally, denoting by $\theta = (A, \Gamma, c, s^2, \beta)$ the vector of parameters of the model, we maximize

$$\log\hat{L}\left(\theta\right) = \log\hat{L}_1\left(\theta\right) + \log\hat{L}_2\left(\theta\right) + \log\hat{L}_3\left(\theta\right),$$

based on the observed $\{y_j, x_j, t_j, Z_j\}$.

[23] These direct effects are not reflected in the matching or wages because if a team is more productive with every potential partner, that does not affect the relative probability of matching with one or how much they would pay different partners.

4.3.2 Identification Without Wages

If wages are not observed, matching patterns generally only identify the sum of the productivity and amenity of a given match, $\frac{\alpha_{y,x}+\gamma_{y,x}}{\sigma^W+\sigma^F}$. However, if there is sufficient variation in tax rates – either separate markets with different tax rates but the same match surplus function, or agents in a single market facing different tax rates – and tax rates are separately observed then $\frac{\alpha_{y,x}}{\sigma^W}$ and $\frac{\gamma_{y,x}}{\sigma^F}$ can be separately identified. The intuition is that if agents of type y_f are more likely to match with agents of type x_w in areas with high taxes, then more of the surplus from that match must be in $\frac{\alpha_{y,x}}{\sigma^W}$, which gets more weight under higher taxes. If agents are more likely to be matched under low taxes, then more of the surplus is in $\frac{\gamma_{y,x}}{\sigma^F}$. We implicitly use this variation in working with the alternative utilities $\tilde{\gamma}_{,}$ and $\tilde{\alpha}_{,}$, but we do not try to identify the utilities without the wage data.

5 Application

As a proof of concept, we analyze the impact of taxation on the matching market of head coaches and college football teams in the Division I Football Bowl Subdivision (FBS) of the National Collegiate Athletic Association (NCAA) in the United States. Football activities in this division generate about $3B per year in revenues for the participating colleges. The market for football coaches is well-suited to our model for at least three reasons. First, college football coaching is an industry where we think preference heterogeneity is potentially quite important. Coaches have different styles of coaching, which can work differently for different teams. Second, there is good data available: not only do we have data on coaches' salaries and characteristics, and colleges' characteristics but team performance gives us a natural measure of productivity, which is not available in most employment datasets. Third, the average yearly salary of head coaches in the Division I FBS is about $1.8M, which is well above the cutoff for the top marginal tax bracket in all states. A linear approximation of taxation of coaches' salaries is therefore justified. However, this market features some limitations. In particular, contracts are often complex, involving bonus incentives and multiple periods. For these reasons, we consider total pay, which includes bonuses, as the compensation of coaches.[24] Our model then predicts the level of the compensation of coaches in equilibrium but is agnostic about how this compensation should be paid.

[24]Many coaches' contracts also specify buyout clauses. If buyout frictions are uncorrelated with observable attributes, these frictions should be captured in the idiosyncratic shocks and hence not affect our results; in contrast, if buyout frictions were correlated with observable attributes of CEOs and firms, the mapping from choice probabilities to parameters (amenities α and productivity γ) would be affected. Unfortunately, estimating any resulting bias would require a dynamic matching model, which is beyond the scope of this paper.

Our data allow us to separately estimate coaches' preferences for job amenities and teams' productivity; the estimates can then be used to simulate the market under alternative tax policies and compare the deadweight loss of taxation, accounting for the matching market, to the deadweight loss estimated ignoring the preference heterogeneity.

5.1 Data

We use data from the 2013 NCAA FBS season, the most recent season for which complete information about teams and coaches was available at the time of the analysis. We hence study the matching market of coaches to teams at the start of the 2013 season. We use the 115 schools, out of 126 in the FBS Division, for which wages are publicly available.[25] The 126 teams are grouped into 10 conferences and include 3 independent teams (Army, Air Force, and Navy). Although the conferences are geographically organized, there are typically schools from different states in each conference.

5.1.1 Sources

We combine data from several sources. The "Coaches Salaries of the Division I FBS" database (USA Today, 2013) provides the salary of each head coach for the 2013 season, his age at the start of the season and his alma mater. Using data from Wikipedia (2013), we construct measures of coaches' experience and ability, proxied respectively by their numbers of games as head coach and their shares of games won, both measured at the start of the 2013 season.

For colleges' characteristics, we get data for each college from the Department of Education (Office of Postsecondary Education, 2013) on football operating expenses and revenues during the 2013 season.[26] Operating expenses are defined as "lodging, meals, transportation, uniforms, and equipment for coaches, team members, support staff." Football revenues include "all revenues attributable to intercollegiate athletic activities,"[27] but a large share of them come from the colleges' share of conference-television network agreements. These TV rights contracts are mostly negotiated at the conference level, sealed for a multi-season period and, each season, the corresponding amounts are distributed to colleges within the

[25] Twenty of the schools are either private or public schools exempted by state law from releasing salary data; nevertheless, as nine of those schools voluntarily released salaries information, salary data are missing for only eleven schools.

[26] We hence assume that these variables are good proxies of a team's attributes that matter for the match value of that team at the start of the season.

[27] Revenues are defined as "all revenues attributable to intercollegiate athletic activities; this includes revenues from appearance guarantees and options, contributions from alumni and others, institutional royalties, signage and other sponsorships, sport camps, state or other government support, student activity fees, ticket and luxury box sales, and any other revenues attributable to intercollegiate athletic activities."

conference according to a pre-determined sharing rule. Therefore, a college's football revenues in any given season do not directly result from the performance of that team in that season but, as we reason, are indicative of the potential for performance and attractiveness of the team for a coaching job.

Finally, we also collect a few measures of team performance:

- the numbers of wins, losses and ties in the 2013 season (National Collegiate Athletics Association, 2013);
- the Football Power Index (FPI) which measures the strength of the team as the season progresses, hence allowing us to measure the change between the start and the end of the season (ESPN, 2013);[28] and
- the Football Recruiting Team Ranking (FRTR) which indicates how well a college has been able to recruit, principally under the impulsion of the head coach (247Sports, 2013).

Lastly, we collect information about the federal and states' income and payroll tax rates incurred by both employees and employers from Tax Foundation (2013).

5.1.2 Summary

Table (1) presents summary statistics for the variables of interest. At the start of the 2013 season, the average head coach was 51 years old, with 95 games of experience; the average ability amounted to one win every two games. In the 2013 season, the average yearly salary of head coaches was approximating \$1.8M, although pay differentials across coaches were quite large as indicated by the relatively large standard deviation (i.e. \$1.3M). The average yearly revenues from football activities in the division in 2013 were about \$28M while operating expenditures averaged roughly \$3.4M.

5.2 Estimates of Job Amenities and Productivity

We apply the estimation strategy described in Section 4.3 to estimate the parameters of the model. Considering age, experience and ability and football revenues and operating expenses as the relevant attributes of coaches and teams, respectively, we obtain estimates of the direct effects of coaches' and teams' attributes, as well as their interaction on both job amenities and productivity. Note that all variables are standardized such that each coefficient can be

[28] According to ESPN (2013), the FPI "is a measure of team strength that is meant to be the best predictor of a team's performance going forward for the rest of the season. FPI represents how many points above or below average a team is."

interpreted as the effect of a one standard deviation change in the associated variable on the productivity or amenity of the match, measured in millions of dollars. Using information about coaches' alma maters allows us to estimate the job amenity and productivity effects of coaching one's alma mater team. We set $\sigma^W = 0.385$ and $\sigma^F = 0.01$ for the remainder of the analysis; these values were selected by performing a grid search and retaining the pair of parameters yielding the largest maximum likelihood value while limiting negative wage predictions to a maximum mass of 5% for each team.[29]

Our estimation strategy fits observed wages well ($R^2 = 0.78$) and fits the performance ranking moderately well (McFadden pseudo-$R^2 = 0.06$).[30] Table (2) presents estimates for the direct and interacted effects of school and coach characteristics on job amenities and productivity. For job amenities, we find, unsurprisingly, that coaches have a substantial willingness-to-pay – about one standard deviation of wages – for coaching their alma maters (\$1.12M, significant at 1%). Coaches also prefer heading teams with lower football revenues. A one standard deviation increase in football revenues decreases the average amount coaches like working for a team by \$630K. However, this effect is mitigated by coaches' experience: an one standard deviation increase in a coach's experience attenuates this effect by \$160K (significant at 5%). We interpret the effect of football revenues as reflecting the impact of higher pressure on the coach's shoulders, since football revenues are mostly resulting from higher media exposure (higher TV rights). Our results also show that abler and younger head coaches prefer teams with larger operating expenses. Head coaches with ability one standard deviation above average (one standard deviation younger than average) derive an additional job enjoyment worth \$260K (resp. \$210K) when heading a team with operating expenses one standard deviation above average.

Regarding teams' productivity, as presented in the lower part of Table (2), two important results are worth noting. First, coaching one's alma mater increases productivity, but the effect is only \$180K (significant at 5%), much smaller than the effect on job amenities. Second, after controlling for the operating budget, football revenues only increase productivity for high-ability coaches. With an average coach, a one standard deviation increase in football revenues decreases productivity by \$180K, but with a coach whose ability is one standard deviation above the mean, the same increase in football revenues increases productivity by

[29] For a wide range of values of σ^W and σ^F, the model predicts observed wages very well with an $R^2 > 0.7$ and no negative wages for observed coach-team pairs. However, negative wages do occur for other pairs, i.e. out-of-sample coach-team pairs, depending on the values of σ^W and σ^F.

[30] As noted in McFadden (1977), p. 35, values of McFadden pseudo-R^2 between .2 and .4 represent an excellent fit. Note also that we account for the fact that the ranking of teams at the bottom of the performance distribution is noisy and hence difficult to predict. We do so by modifying the likelihood function such that only the ranking of the best 15 teams is considered informative, hence assuming the ranking beyond the 15^{th} place is as good as random (see, e.g., Fok et al. (2012)).

$\$60\text{K} = \$240\text{K} - \$180\text{K}$. Our result is in line with our interpretation of football revenues as reflecting pressure put on the team and the coach, suggesting that higher pressure can only be dealt with by high ability coaches.

5.3 Simulations

We can use our estimated parameters to calculate the equilibrium match probabilities and wages, and hence social welfare, under alternative tax policies.[31] As noted in section 4.3, given parameters (A, Γ), the equilibrium matching probabilities, and hence the equilibrium wages, can be obtained up to a normalizing constant c. The normalization arises because unmatched agents (unemployed workers and inactive firms) are typically not observed in the data. For the observed market, the associated normalizing constant is estimated using observed wages, but when performing simulations the constant for the counterfactual conditions (tax rates) needs to be specified by the analyst. Note, however, that the change in the normalizing constant depends on how agents split the surplus gained or lost from the change in tax. If coaches have the bargaining power (are on the short side of the market), then the firms' outside options determine the overall wage level in the market, meaning that wages for a given match will not change a lot when the tax rate changes, though the probability of matchings with high or low wages will change. Conversely, if firms have most of the bargaining power, then coaches' outside options determine the overall wage levels. As taxes decrease, there will be a corresponding decrease in the overall wage level to keep the post-tax wages mostly unchanged.

Unfortunately we do not observe unmatched agents and therefore cannot estimate the relative bargaining power of the two sides of the market. Instead, we consider four different alternative assumptions about how the bargaining power is distributed between the two sides of the market and derive – for each of these alternatives – the associated equilibrium wages at each tax rate. The first two alternatives correspond to the extreme points of the distribution of bargaining power, met when either side of the market has no power (is on the long side of the market). The other two alternatives are intermediate points. The first intermediate alternative simply corresponds to an equal distribution of power resulting in a 50-50 split of surplus. For the second intermediate case, we note that it is less "efficient" from the perspective of a worker-firm pair to shift surplus to workers because doing so increases the tax burden; we calculate wages based on an efficiency-weighted split of the surplus between

[31] Because sports teams compete with each other, improving the quality all teams' coaching may not change the balance among teams substantially. Nevertheless, sports coaching is not zero-sum, in welfare terms: improving all the coach–team matches would improve game quality, resulting in higher revenues (see the related discussion by Fréchette et al. (2007)).

workers and firms.[32]

For each tax rate, the equilibrium match probabilities and wages allow one to calculate welfare and also tax revenue. As we vary the tax rate, we calculate both the mechanical change in revenue,

$$\sum_y \sum_x \mu_{yx} t_{y,x} \Delta\tau,$$

and the change in revenue resulting from changes in the match patterns and wages,

$$\sum_y \sum_x \tau \left(\mu_{yx} t_{y,x} - \mu'_{yx} t'_{y,x} \right). \tag{5.1}$$

In models of non-matching labor markets, (5.1) is a measure of deadweight loss (DWL) from the incremental tax increase (see Feldstein (1999) and the discussion in Section 3.4). We compare the wage-based measure of DWL from (5.1) to the actual DWL:

$$\sum_y \sum_x \left(\mu_{yx} - \mu'_{yx} \right) \left(\gamma_{y,x} + \alpha_{x,y} \right). \tag{5.2}$$

Note that since the Feldstein measure of DWL in (5.1) depends on wages and wages are obtained using four alternative distributions of bargaining power, we therefore have four different measures of DWL to compare to the true value (calculated using (5.2)).

5.3.1 Federal Top Tax Rate and State Taxes

We first study the effect of varying the federal top tax rate on welfare over the range (0, .5) which contains the 2013 value, i.e. 0.4195.[33] The results are presented in Figures 4 and 5. Figure 4 shows the true DWL and three measures of the DWL estimated from wage changes. For the three alternative distributions of bargaining power shown – the measure obtained when firms have all the bargaining power, the measure obtained when firms and workers split the surplus 50-50, and the measure obtained under the efficiency-weighted split of surplus – wages are increasing when taxes increase, causing the DWL estimated from wage changes to be of the wrong sign. The wage-based measures of DWL ignore changes in firm surplus. When firms' have bargaining power, their surplus is decreasing in the tax rate; That decrease is not captured in the the three wage-based measures of DWL shown in Figure 4.

As shown in Figure 5a, the DWL estimated from wage changes under the assumption that coaches have all of the bargaining power is a more reasonable estimate of the true DWL.

[32] If t^1 is the transfer that keeps the average coaches' payoff fixed (teams have all the bargaining power) and t^2 is the transfer that keeps the average teams' payoff fixed, the efficiency weighted transfer is $t = \frac{t^1(1-\tau)^2 + t^2}{(1-\tau)^2+1}$.

[33] The 41.95% rate includes federal payroll taxes.

In that case, the wage for a given match does not change much with the tax rate since it is pinned down by firms' outside options. As a result, wage changes largely reflect workers moving to jobs at which they are less productive.[34] However, the amenities at those jobs are higher, which is not reflected in the wage; thus the estimate of DWL from wage changes generally over-estimates the true DWL from a tax increase. The relationship between the two is different for high tax rates because of interactions with state taxes. Figure 5b graphs the difference between the true and the DWL estimated from wage changes as a fraction of the true value. With state taxes set to 0, the wage-based estimate of DWL is always more negative than the true value. In contrast, with observed state taxes, the relationship flips for high tax rates.

We next study the effect of varying state taxes on welfare. We do so by varying the average level of state taxes both with taxes varying in proportion to their observed levels and with tax equalized across states. The potential for substitution across states raises the question of the welfare and revenue effects of equalizing state taxes. To answer this question, Figure 6 plots welfare and revenue relative to the baseline levels for equalized and unequalized taxes. We vary the on equalized tax rates from 0 to 100% of their observed values; the equalized tax rate varies from zero to 4.8%, which is the average rate across schools in the data. The line for equalized tax rates (dashed line) is down and to the left of the line for tax rates proportional to their observed levels (sold line). Equalizing tax rates across states for the observed average level of taxes both lowers revenue and lowers welfare.

5.3.2 Thresholds and Inframarginal Rates

We next use simulations to look at the effect of varying the federal tax bracket thresholds and inframarginal tax rates on welfare; these exercises require that we account for the fact that the observed federal tax schedule is piecewise linear rather than linear. In Appendix A.4 we describe how the convexity of a piecewise linear, progressive tax schedule allows a simple extension of the linear model. Intuitively, one can think of the $\langle$transfer, post-tax transfer$\rangle$ pairs that are feasible under a piecewise linear tax as the intersection of the pairs that are feasible under each of the underlying linear taxes, albeit after adjusting the utilities to account for the fact that the linear schedules for the higher brackets do not intersect the origin. As a result, rather than solving for the individual expected utilities ($\{b\}$ and $\{a\}$) such that each agent's match probabilities sum to 1, we find the expected utilities such that the minimum probability across each tax bracket sums to 1.

We first study the effect on welfare and revenue of varying the inframarginal tax rate with

[34] If we do not account for the change in match probabilities, increasing the tax causes a very slight increase in average wage.

piecewise linear taxation. When changing the inframarginal tax rate, whether the coaches or teams are the residual claimants to additional surplus affects not only the revenue, but also the welfare. Lowering the inframarginal tax rate increases post-tax wages, so when coaches' utility is kept constant, there must be an accompanying decrease in the overall wage level. The overall decrease in wages lowers the marginal tax rates for some coaches and teams leading to an increase in efficiency. When teams' utility is kept constant wages for a given match do not change systematically or by very much, so there is no secondary effect of overall wage levels on efficiency. As a consequence, welfare is more sharply decreasing in the inframarginal tax rate when coaches utility is held constant than when teams' utility is held constant, as shown in Figure 7a. Since wages are increasing in the tax rate when coaches utility is fixed, revenue also increases more quickly in the tax rate as depicted in Figure 7b. Comparing the scales between Figure 7a and Figure 7b, we note that the differences in social welfare across different inframarginal tax rates are very small relative to the differences in revenue.

Finally, we can also vary the threshold for the top tax bracket.[35] The difference in the tax rate above and below the threshold is 6.6 percentage points. Figure 8 shows welfare (Figure 8a) and revenue (Figure 8b) as the threshold for the top bracket moves between \$200k and \$600k. Again, and for the same reasons, both are more affected by the threshold when coaches' utility is held fixed than when teams' utility is held fixed.

6 Conclusion

This paper investigates the incidence of taxation in matching markets both theoretically and empirically. On the theoretical side, we show through examples and numerical simulations that, in general, the total match value of stable matchings may not vary monotonically as taxation increases. We then show three important positive results for markets that are wage markets under proportional taxation. First, as proportional taxes decrease, the total match value of stable matchings (weakly) increases. Second, while in the absence of taxation, stable matchings maximize the sum of match values of workers and firms with unit weights for workers and firms, with tax rate $\tau > 0$, stable matchings still maximize the sum of match values of workers and firms, but with respective weights 1 and $(1 - \tau)$. Third, workers' match values are weakly increasing in the tax rate whereas firms' match values are weakly decreasing in the tax rate.

Using insights from the second result above, we adopt the Choo and Siow (2006) approach

[35] Since the tax bracket starts at \$400k with a tax rate of 41.95% and the bracket below that starts at \$398k with a tax rate of 37.35%, we remove the second-from-the-top bracket.

on renormalized utilities to account for the linear taxation. Since the renormalization also applies to the random components of the utilities, the distributions of unobserved heterogeneity are fundamentally heteroskedastic as soon as taxation varies with observed types x and y. We therefore adapt the Dupuy and Galichon (2017) maximum likelihood estimator of amenities and productivities when transfers are observed with noise to allow for this heteroskedasticity. Finally, we extend the maximum likelihood estimator to the case in which noisy measures of firm productivity are also observed.

We use our estimation strategy to study the matching market of head coaches and college football teams in the Division I Football Bowl Subdivision (FBS) of the National Collegiate Athletic Association (NCAA) in the United States. We estimate that coaches are willing to give up \$1.12M in order to coach their alma mater team and prefer heading teams with low pressure to perform, although this latter preference is decreasing with coach experience. We also find that teams are \$180K more productive when coached by alma mater coaches; meanwhile, team productivity decreases with the pressure to perform unless the team is coached by a high enough ability coach.

We perform simulations of the impact of (federal and state) tax policies based on our estimates. Results confirm that wage-based measures of DWL, as suggested by Feldstein (1999), over-estimate the true DWL from a tax increase if wages do not respond to the tax, because these measures miss the higher amenities that workers receive from lower wage jobs. Simulations also show that a proportional increase in the average level of state taxes decreases the average workers' amenities because productive coaches substitute towards jobs in states with lower tax rates rather than just jobs with higher amenities. The substitution effect we observe can have important welfare implications as our results show: equalizing tax rates across states for the observed average level of taxes both lowers tax revenue and welfare.

Table 1: Summary statistics of coaches' and teams' attributes

	Mean	Std	Min	MAx
Coaches:				
Age (in years)	50.70	8.50	33.00	74.00
Experience (#games)	95.04	86.98	0.00	389.00
Ability (wins/game)	0.51	0.23	0.00	0.91
Salaries (in $M)	1.77	1.30	0.29	5.55
Coaches at Alma Mater	0.09	0.28	0.00	1.00
Teams:				
Football revenues (in $M)	28.02	23.94	4.11	112.51
Operating expenses (in $M)	3.35	2.30	0.60	15.23
Performance: Principal Component	-0.00	1.35	-3.04	3.65
N= 115				

Note: Salaries refers to *total pay* which includes *school pay* – the base salary paid by the university plus other income paid, or guaranteed, by the university – and *other pay* not guaranteed by the university. On average, school pay represents 99% of total pay. The performance measure is the principal component of a PCA performed on (1) the number of wins in the season, (2) the change in FPI index between start and end of season and (3) the FRTR ranking.

Table 2: Effect of coaches' and teams' attributes on job amenities and productivity (in \$M)

	Main effects	Football revenues (in \$M)	Operating expenses (in \$M)
Job Amenities (Alpha)			
Main effects		-0.63	-0.11
		(0.07)	(0.07)
Age (in years)		-0.09	-0.21
		(0.09)	(0.09)
Experience (# of games)		0.16	0.10
		(0.08)	(0.09)
Ability (wins/game)		0.09	0.26
		(0.10)	(0.09)
Alma Mater	1.12		
	(0.14)		
Productivity (Gamma)			
Main effects		-0.18	0.35
		(0.09)	(0.16)
Age (in years)	-0.04	-0.00	0.16
	(0.05)	(0.07)	(0.09)
Experience (# of games)	0.11	0.00	-0.19
	(0.06)	(0.06)	(0.11)
Ability (wins/game)	0.13	0.24	-0.10
	(0.07)	(0.11)	(0.12)
Alma Mater	0.18		
	(0.09)		
Scaling performance	7.22		
	(3.25)		
Salary constant	8.05		
	(0.06)		

Note: This table reports the estimates of effects the interaction of coach characteristics and team characteristics on team productivity and job amenities, measured in millions of dollars. All covariates except for alma mater are standardized to have a standard deviation of 1. Standard errors, calculated from the Hessian of the likelihood, are in parentheses. Scaling performance refers to the variance of the error in the performance equation (β).

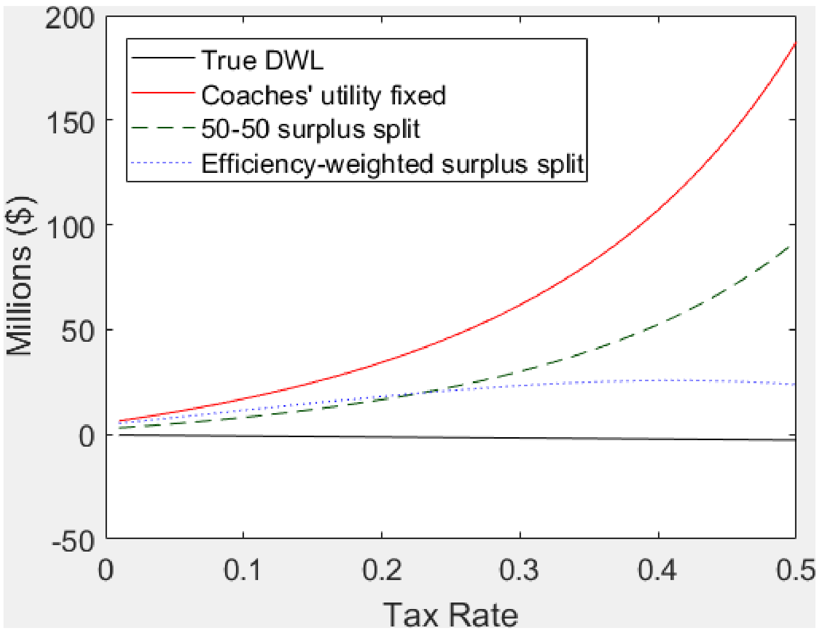
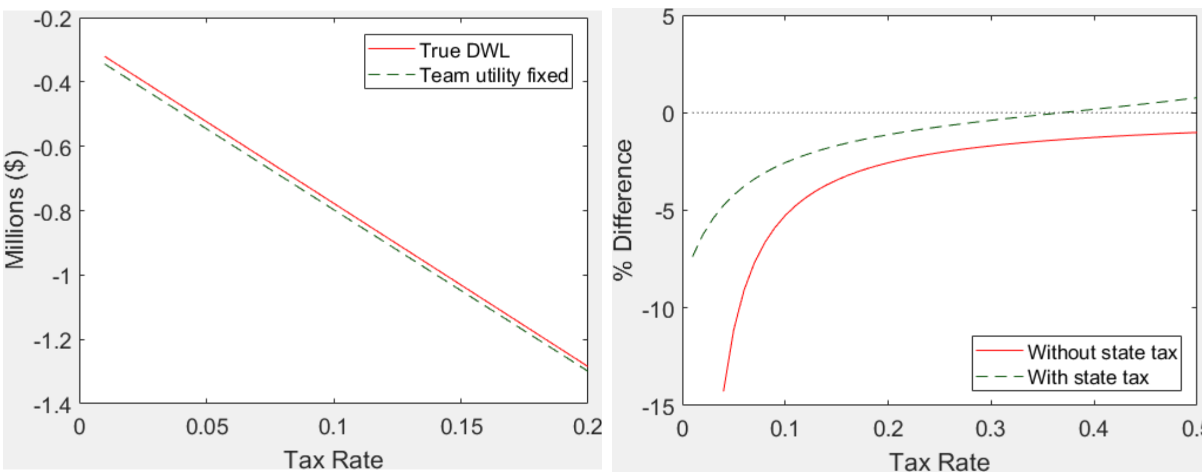


Figure 4: Marginal deadweight loss from a percentage point increase in the federal tax

Note: This graph shows the deadweight loss (DWL), in millions of dollars, from a one percentage point increase in the federal tax rate. The bottom line is the true DWL from the simulation and top three lines are the DWL estimated based on the changes in wages under three different assumptions about how wages adjust to account for the change in surplus from the change in the tax rate. The first setting holds the coaches' utility fixed so any additional surplus goes to the teams; the second splits the surplus 50-50; and the last assumes an efficiency-weighted split of the surplus as described in Section 5.3. All of the wage-based estimates are of the wrong sign because wages increase to compensate workers for loss from the higher tax rate.



(a) True and estimated

(b) % Error in estimate

Figure 5: Marginal deadweight loss from increase in the federal tax

Note: This graph shows the deadweight loss (DWL), in millions of dollars, from a 1 percentage point increase in the federal tax rate. The solid line shows the true DWL from the simulation and the dashed line shows the DWL estimated based on the changes in wages assuming that coaches have all the bargaining power (where wages change very little in response to changes in the tax rate). The wage-based estimate over estimates the DWL because it does not account for the higher amenities workers receive in matchings where they are less productive and paid lower wages.

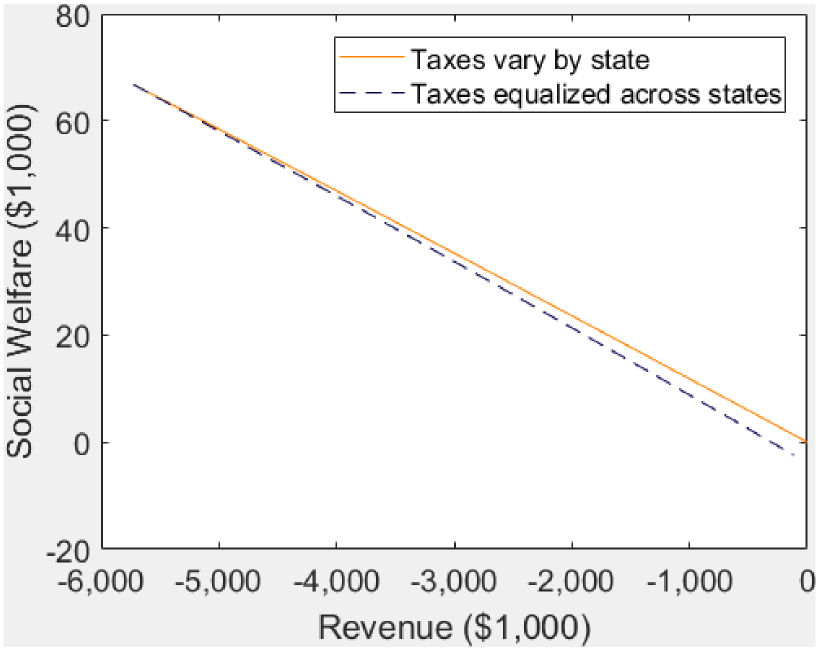
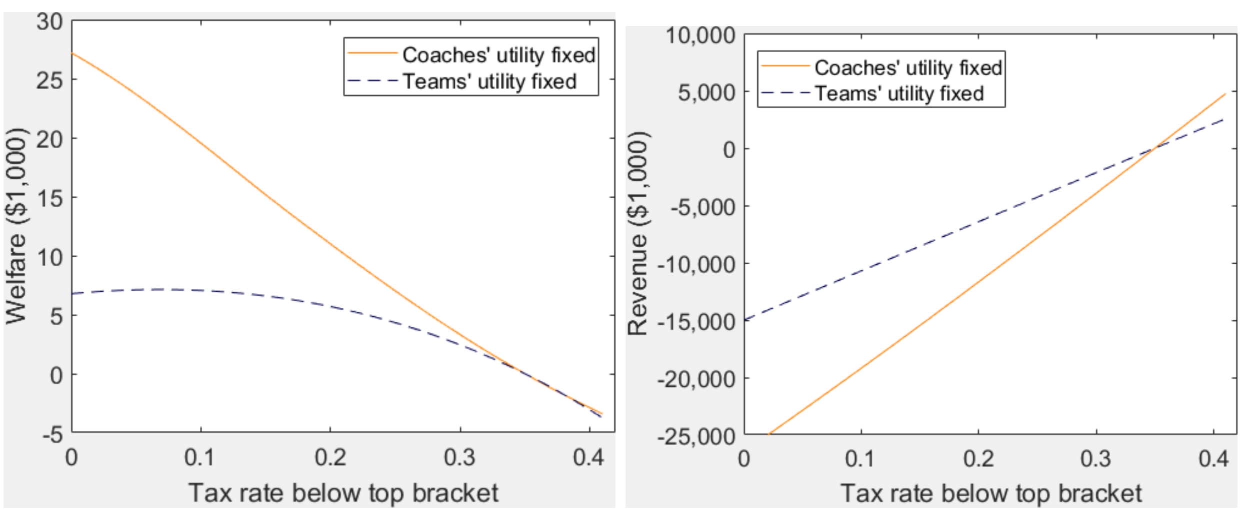


Figure 6: Relationship between welfare and revenue

Note: This graph shows how welfare and revenue co-vary as the state tax rate changes. Both quantities are in thousand of dollars and relative to baseline levels. The mean tax rate is varied from 0 to 4.8%, first keeping taxes proportional to their observed level and then for tax rates equalized across states. For a given revenue, welfare is lower when taxes are equalized across states.

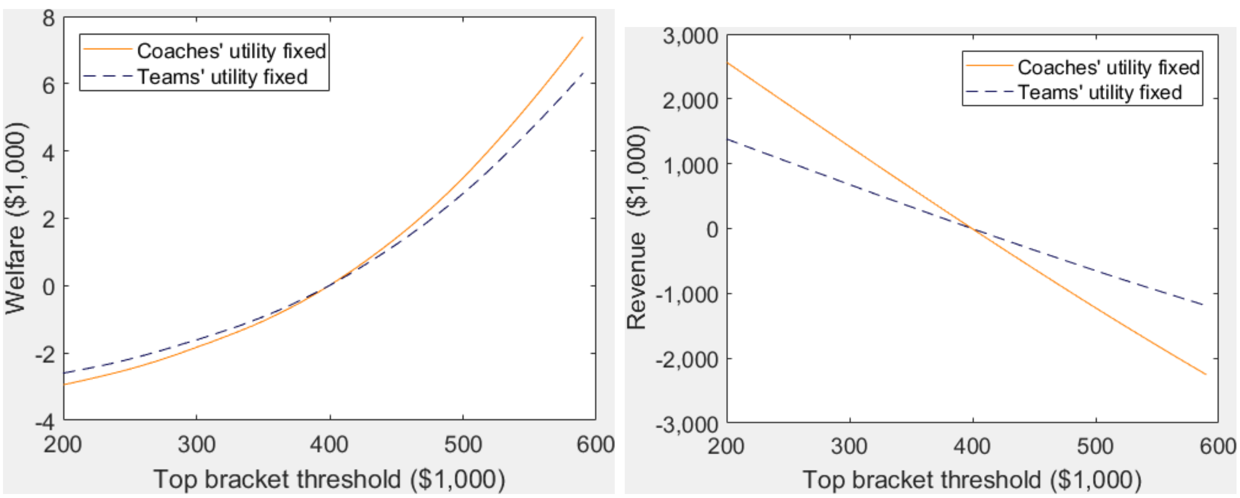


(a) Welfare

(b) Revenue

Figure 7: Effect of the infra-marginal tax rate

Note: These graphs show the effect of changing the inframarginal tax rate on tax revenue and social welfare. Welfare is in thousand of dollars and relative to level with the baseline level with a 35% inframarginal tax rate. The top marginal tax rate is kept constant at .4195. All income below 400k is treated as a single bracket whose tax rate varies from 0 to .4195. The effects on revenue are much larger than those on welfare.



(a) Welfare

(b) Revenue

Figure 8: Effect of the threshold for the top tax bracket

Note: These graphs show the effect of changing the threshold above which the top marginal tax rate of 41.95% applies. Quantities are in thousand of dollars and relative to baseline levels. To make the effect clearer, we drop the second from the top bracket so all income below the threshold and above $183 is taxed at 35.35%. So the change in marginal tax rate at the threshold is 6.6%. The effects on revenue are much larger than those on welfare.

Appendix

A Additional Theoretical Results

A.1 Taxes cannot lead to Pareto improvements

Although Example 1 and the simulations show that total match value of stable matchings may decrease when the tax rate falls, an arrangement that is stable under a tax rate $\hat{\tau}$ must raise the payoff of at least one agent, relative to an arrangement that is stable under a tax rate $\check{\tau} > \hat{\tau}$.

Proposition 2. *Suppose that $[\hat{\mu}; \hat{t}]$ is stable under tax $\hat{\tau}$, and that $[\check{\mu}; \check{t}]$ is stable under tax $\check{\tau}$, with $\check{\tau} > \hat{\tau}$. Then, $[\check{\mu}; \check{t}]$ (under tax $\check{\tau}$) cannot Pareto dominate $[\hat{\mu}; \hat{t}]$ (under tax $\hat{\tau}$).*[36]

To see the intuition behind Proposition 2, we consider the case in which $\check{\tau} = 1$ and choose $\check{t} = 0$: If $[\check{\mu}; \check{t}]$ (under tax $\check{\tau} = 1$) Pareto dominates $[\hat{\mu}; \hat{t}]$ (under tax $\hat{\tau}$), then every firm $f \in F$ (weakly) prefers $\check{\mu}(f)$ to $\hat{\mu}(f)$ with the transfer $\hat{t}_{f,\hat{\mu}(f)}$.[37] But then, because $[\hat{\mu}; \hat{t}]$ is stable under tax $\hat{\tau}$,

$$\gamma_{f,\check{\mu}(f)} \overbrace{\geq}^{\text{Pareto}} \gamma_{f,\hat{\mu}(f)} - \hat{t}_{f,\hat{\mu}(f)} \overbrace{\geq}^{\text{Stability}} \gamma_{f,\check{\mu}(f)} - \hat{t}_{f,\check{\mu}(f)},$$

so every f must be offering a weakly positive transfer to $\check{\mu}(f)$ under $\hat{t}$ (that is, $\hat{t}_{f,\check{\mu}(f)} \geq 0$). An analogous argument shows that each worker $w \in W$ must be offering a weakly positive transfer to $\check{\mu}(w)$ under $\hat{t}$ (that is, $\xi_{\hat{\tau}}(\hat{t}_{\check{\mu}(w),w}) \leq 0$). Moreover, Pareto dominance implies that at least one firm or worker must be paying a *strictly* positive transfer. But then, that agent must pay a strictly positive transfer and receive a weakly positive transfer – impossible.

A.2 The Effect of Very Small Taxes

Unlike in non-matching models of taxation, in our setting there is always a nonzero tax that does not generate distortions. To find the minimum tax that generates a distoriton, let $\hat{\mu}$ be an efficient matching. Our results show that if $\check{\mu}$ is stable under $\check{\tau}$, then[38]

$$\check{\tau} \geq \frac{\mathfrak{M}(\hat{\mu}) - \mathfrak{M}(\check{\mu})}{\sum_{f \in F} (\gamma_{f,\hat{\mu}(f)} - \gamma_{f,\check{\mu}(f)})}. \tag{A.1}$$

[36] We say that an arrangement $[\check{\mu}; \check{t}]$ under tax $\check{\tau}$ *Pareto dominates* arrangement $[\hat{\mu}; \hat{t}]$ under tax $\hat{\tau}$ if

$$\begin{aligned} \gamma_{f,\check{\mu}(f)} - \check{t}_{f,\check{\mu}(f)} &\geq \gamma_{f,\hat{\mu}(f)} - \hat{t}_{f,\hat{\mu}(f)} & \forall f \in F, \\ \alpha_{\check{\mu}(w),w} + \xi_{\check{\tau}}(\check{t}_{\check{\mu}(w),w}) &\geq \alpha_{\hat{\mu}(w),w} + \xi_{\hat{\tau}}(\hat{t}_{\hat{\mu}(w),w}) & \forall w \in W, \end{aligned}$$

with strict inequality for some $i \in F \cup W$.

[37] Note that under tax $\check{\tau} = 1$, an arrangement with transfers of 0 among match partners Pareto dominates any other arrangement with the same matching. Thus, the transfers between match partners under $[\check{\mu}; \check{t}]$ can be assumed to be 0. Then, the comparison between $[\check{\mu}; \check{t}]$ (under tax $\check{\tau} = 1$) and $[\hat{\mu}; \hat{t}]$ (under tax $\hat{\tau}$) amounts to a comparison of agents' match values under $\check{\mu}$ and their total utilities under $[\hat{\mu}; \hat{t}]$.

[38] See Section B.4.

For any inefficient matching $\breve{\mu}$, there is a strictly positive minimum tax $\underline{\tau}(\breve{\mu})$ at which $\breve{\mu}$ could possibly be stable. Since there are finitely many possible matchings, we can just take the minimum of the threshold $\underline{\tau}$ across inefficient matchings,

$$\underline{\tau}^* = \min_{\{\mu:\mathfrak{M}(\mu)<\mathfrak{M}(\hat{\mu})\}} \left[\underline{\tau}(\mu)\right].$$

For $\tau < \underline{\tau}^*$ only an efficient matching can be stable.[39]

A.3 Multiple Matchings Stable at a Given Tax Rate

Hatfield et al. (2013) have shown that in markets with perfect transfers, if two different matchings are stable, then they can be supported by the same transfer vector and all agents are indifferent between the resulting arrangements; these results carry over to the case of taxation (in wage markets).

Proposition 3. *If two matchings $\breve{\mu}$ and $\hat{\mu}$ are both stable in a wage market at tax rate τ, then:*

1. *Any transfer vector that supports $\breve{\mu}$ also supports $\hat{\mu}$, and vice versa.*
2. *For any transfer vector t supporting $\breve{\mu}$ and $\hat{\mu}$, all agents are indifferent between $[\breve{\mu}, t]$ and $[\hat{\mu}, t]$ (i.e., $u_i([\breve{\mu}, t]) = u_i([\hat{\mu}, t])$).*
3. *For any transfer vector t supporting $\breve{\mu}$ and $\hat{\mu}$, the difference in revenue under $\breve{\mu}$ and $\hat{\mu}$ equals the difference in total match value between $\breve{\mu}$ and $\hat{\mu}$.*

Proof. For any stable matchings $\hat{\mu}$ and $\breve{\mu}$, and a transfer vector t that supports one of the stable matchings—say, $\hat{\mu}$—we can renormalize worker match utilities $\{\tilde{\alpha}_{f,w}\} = \frac{1}{1-\tau}\{\alpha_{f,w}\}$ to obtain a market in which the results of Hatfield et al. (2013) apply. Indeed, by (Hatfield et al., 2013, Theorem 3), we must have

$$\gamma_{f,\breve{\mu}(f)} - t_{f,\breve{\mu}(f)} = \gamma_{f,\hat{\mu}(f)} - t_{f,\hat{\mu}(f)} \qquad \forall f, \tag{A.2}$$

$$\frac{1}{1-\tau}\alpha_{\breve{\mu}(w),w} + t_{\breve{\mu}(w),w} = \frac{1}{1-\tau}\alpha_{\hat{\mu}(w),w} + t_{\hat{\mu}(w),w} \qquad \forall w. \tag{A.3}$$

If we multiply both sides of (A.3) by $1-\tau$, we obtain

$$\alpha_{\breve{\mu}(w),w} + (1-\tau)t_{\breve{\mu}(w),w} = \alpha_{\hat{\mu}(w),w} + (1-\tau)t_{\hat{\mu}(w),w} \qquad \forall w. \tag{A.4}$$

We conclude from (A.2) and (A.4) that t must support $\breve{\mu}$ as well as $\hat{\mu}$. Moreover, under transfer vector t, every agent is indifferent between $\hat{\mu}$ and $\breve{\mu}$.

Multiplying (A.3) by $(1-\tau)$ and summing across agents gives

$$\mathfrak{M}(\breve{\mu}) - \mathfrak{M}(\hat{\mu}) = \tau \sum_{f\in F} t_{f,\breve{\mu}(f)} - \tau \sum_{f\in F} t_{f,\hat{\mu}(f)}. \qquad \square$$

[39] One caveat is that if there are multiple efficient matchings (all of which are stable when $\tau = 0$), some of them may *not* be stable in the limit as $\tau \to 0$.

Unfortunately the third statement in Proposition 3—that changes in revenue reflect changes in value—is very limited. As the tax rate changes, transfers will change even when the underlying matching does not change (so there is no change in total match value). Also, even at a tax rate under which multiple matchings are stable, there may be multiple supporting transfer vectors $\hat{t}$ and $\check{t}$, and the revenue comparison between $[\hat{\mu}, \hat{t}]$ and $[\check{\mu}, \check{t}]$ does not tell us anything about the difference in total match value between $\hat{\mu}$ and $\check{\mu}$.

Results of Kelso and Crawford (1982) and Hatfield and Milgrom (2005) imply that for any fixed τ, if there are multiple stable arrangements, then workers' and firms' interests are opposed: If all firms prefer $[\mu; t]$ to $[\hat{\mu}; \hat{t}]$, then all workers prefer $[\hat{\mu}; \hat{t}]$ to $[\mu; t]$. Moreover, there exists a firm-optimal (worker-pessimal) stable arrangement that the firms weakly prefer to all other stable arrangements and a worker-optimal (firm-pessimal) stable arrangement that all workers weakly prefer.

Proposition 4. *In a wage market with proportional taxation, if two distinct matchings $\check{\mu}$ and $\hat{\mu}$ are both stable under tax τ, then*

$$\sum_{w \in W} \left(\alpha_{\check{\mu}(w),w} - \alpha_{\hat{\mu}(w),w}\right) = (1-\tau) \sum_{f \in F} \left(\gamma_{f,\hat{\mu}(f)} - \gamma_{f,\check{\mu}(f)}\right). \tag{A.5}$$

Thus, if the firms are not indifferent in aggregate between $\check{\mu}$ and $\hat{\mu}$, then the only tax rate τ under which both $\check{\mu}$ and $\hat{\mu}$ can be stable is

$$\tau = 1 + \frac{\sum_{w \in W} \left(\alpha_{\check{\mu}(w),w} - \alpha_{\hat{\mu}(w),w}\right)}{\sum_{f \in F} \left(\gamma_{f,\check{\mu}(f)} - \gamma_{f,\hat{\mu}(f)}\right)}. \tag{A.6}$$

For τ as defined in Equation (A.6) to be less than 1, the fraction on the right-hand side must be negative, so that that firms and workers in aggregate disagree about which matching they prefer. In wage markets with proportional taxation, where there is generically a unique stable matching, this opposition of interests carries over to the set of supporting transfer vectors.

In order for there to be multiple values of τ at which two given matchings are both stable, both firms and (following (A.5)) workers must be indifferent between those two matchings.

Corollary 1. *In a wage market with proportional taxation, if there is more than one tax under which two distinct matchings $\check{\mu}$ and $\hat{\mu}$ both are stable, then $\mathfrak{M}(\check{\mu}) = \mathfrak{M}(\hat{\mu})$.*

Corollary 1 implies that for generic match values, there is at most one value of τ at which two matchings $\check{\mu}$ and $\hat{\mu}$ are both stable; since there are finitely many matchings, there is a unique stable matching under almost every tax τ.

A.4 Nonlinear Taxes

The wage schedule is in fact progressive, and the worker's income is a concave function of the nominal transfer. We do not have general theoretical results for non-linear taxes, but we can adapt the econometric framework we use to calculate the market equilibrium for the case of piecewise linear taxes.

Assume that the tax rates on workers are $\tau_x^{W,1} < \tau_x^{W,2} < ... < \tau_x^{W,K}$, where $\tau_x^{W,1}$ and $\tau_x^{W,K}$ are respectively the lowest and the top tax rate. Tax rate $\tau_x^{W,k}$ applies to the income above t_x^k, where $t_x^1 = 0$, so that if $\xi_{xy}(t_x^k)$ is the post-tax income of a worker of type x earning t_x^k and working for a firm of type y, then

$$\xi_{xy}(t_x^{k+1}) = \xi_{xy}(t_x^k) + \left(1 - \tau_y^W\right)\left(1 - \tau_x^{W,k}\right)\left(t_x^{k+1} - t_x^k\right).$$

More generally, the post-tax income of a worker of type x working for a firm of type y and earning t_x^k is

$$\xi_{xy}(t) = \min_{k\in\{1,\dots,K\}} \left\{\xi_{xy}(t_x^k) + \left(1 - \tau_y^W\right)\left(1 - \tau_x^{W,k}\right)\left(t - t_x^k\right)\right\}.$$

Hence, the systematic utility of a worker x working for a firm of type y with pre-tax income t is

$$\mathrm{u}_{xy} = \min_{k\in\{1,\dots,K\}} \left\{\alpha_{x,y}^k + \left(1 - \tau_y^W\right)\left(1 - \tau_x^{W,k}\right) t\right\}, \tag{A.7}$$

where $\alpha_{xy}^1 = \alpha_{xy}$, and $\alpha_{x,y}^{k+1} = \alpha_{x,y}^k + \left(1 - \tau_y^W\right)\left(\tau_x^{W,k+1} - \tau_x^{W,k}\right) t_x^{k+1}$. On the firm's side, the systematic utility is still

$$\mathrm{v}_{yx} = \gamma_{y,x} - \left(1 + \tau_y^F\right) t. \tag{A.8}$$

Substituting out t from equations (A.7) and (A.8), we obtain

$$\mathrm{u}_{xy} = \min_{k\in\{1,\dots,K\}} \left\{\alpha_{x,y}^k + \left(1 - \tau_y^W\right)\left(1 - \tau_x^{W,k}\right)\left(\frac{\gamma_{y,x} - \mathrm{v}_{yx}}{1 + \tau_y^F}\right)\right\},$$

which can be rewritten as

$$\min_{k\in\{1,\dots,K\}} \left\{\alpha_{x,y}^k - \mathrm{u}_{xy} + \left(1 - \tau_y^W\right)\left(1 - \tau_x^{W,k}\right)\left(\frac{\gamma_{y,x} - \mathrm{v}_{yx}}{1 + \tau_y^F}\right)\right\} = 0.$$

So, letting $\lambda_x^{W,k} = \frac{1}{1-\tau_x^{W,k}}$ (and $\lambda_y^F = \frac{1-\tau_y^W}{1+\tau_y^F}$, as before), we have

$$\min_{k\in\{1,\dots,K\}} \left\{\alpha_{x,y}^k - \mathrm{u}_{xy} + \frac{\lambda_y^F}{\lambda_x^{W,k}}\left(\gamma_{y,x} - \mathrm{v}_{yx}\right)\right\} = 0,$$

which is equivalent to

$$\min_{k\in\{1,\dots,K\}} \left\{\frac{\lambda_x^{W,k}(\alpha_{x,y}^k - \mathrm{u}_{xy}) + \lambda_y^F\left(\gamma_{y,x} - \mathrm{v}_{yx}\right)}{\sigma^W\lambda_x^{W,k} + \sigma^F\lambda_y^F}\right\} = 0. \tag{A.9}$$

By the log-odds formula, we have

$$\sigma^W \ln \frac{\mu_{xy}}{\mu_{x0}} = \mathrm{u}_{xy}$$
$$\sigma^F \ln \frac{\mu_{xy}}{\mu_{0y}} = \mathrm{v}_{xy},$$

where we adjust μ_{x0} and μ_{0y} so that each agent's match probabilities sum to 1. After plugging the preceding formulas into (A.9), we find that

$$\min_{k\in\{1,\ldots,K\}} \left\{ \frac{\lambda_x^{W,k}(\alpha_{x,y}^k + \sigma^W \ln \frac{\mu_{x0}}{\mu_{xy}}) + \lambda_y^F \left(\gamma_{y,x} + \sigma^F \ln \frac{\mu_{0y}}{\mu_{xy}}\right)}{\sigma^W \lambda_x^{W,k} + \sigma^F \lambda_y^F} \right\} = 0,$$

which implies that

$$\ln \mu_{xy} = \min_{k\in\{1,\ldots,K\}} \left\{ \frac{\lambda_x^{W,k}(\alpha_{x,y}^k + \sigma^W \ln \mu_{x0}) + \lambda_y^F \left(\gamma_{y,x} + \sigma^F \ln \mu_{0y}\right)}{\sigma^W \lambda_x^{W,k} + \sigma^F \lambda_y^F} \right\},$$

which yields

$$\mu_{xy} = \min_{k\in\{1,\ldots,K\}} M_{xy}^k \left(\mu_{x0}, \mu_{0y}\right),$$

where

$$M_{xy}^k \left(\mu_{x0}, \mu_{0y}\right) = \mu_{x0}^{\frac{\sigma^W \lambda_x^{W,k}}{\sigma^W \lambda_x^{W,k} + \sigma^F \lambda_y^F}} \mu_{0y}^{\frac{\sigma^F \lambda_y^F}{\sigma^W \lambda_x^{W,k} + \sigma^F \lambda_y^F}} e^{\frac{\lambda_x^{W,k} \alpha_{x,y}^k + \lambda_y^F \gamma_{y,x}}{\sigma^W \lambda_x^{W,k} + \sigma^F \lambda_y^F}}.$$

We can use the preceding observations to solve for the equilibrium μ_{x0}, μ_{0y} such that the probabilities for each x and each y sum to 1.

B Proofs of Results in Section 2 and Appendix A

B.1 Existence of Stable Arrangements & Equivalence with the Core and Competitive Equilibria

In this section, we use results from the literature on matching with contracts to show the existence of stable arrangements in our framework. For a given transfer vector t, the *demand of firm $f \in F$*, denoted $\mathcal{D}^f(t)$, is

$$\mathcal{D}^f(t) \equiv \arg\max_{D\subseteq W} \left\{\gamma_{f,D} - t_{f,D}\right\}.$$

Definition 3 (Kelso and Crawford (1982))**.** *The preferences of firm $f \in F$ are* substitutable *if for any transfer vectors t and $\check{t}$ with $\check{t} \geq t$, there exists, for each $D \in \mathcal{D}^f(t)$, some $\check{D} \in \mathcal{D}^f(\check{t})$ such that*

$$\check{D} \supseteq \{w \in D : t_{f,w} = \check{t}_{f,w}\}.$$

That is, the preferences of $f \in F$ are substitutable if an increase in the "prices" of some workers cannot decrease demand for the workers whose prices remain unchanged.[40]

Theorem 2 of Kelso and Crawford (1982) shows that under the assumption that all firms' preferences are substitutable, there is an arrangement $[\mu; t]$ that is in the *strict core*, in the sense that:[41]

- Each agent (weakly) prefers his assigned match partner(s) (with the corresponding transfer(s)) to being unmatched, that is,

$$u_i([\mu; t]) \geq 0 \qquad\qquad \forall i \in F \cup W.$$

- There does not exist a firm $f \in F$, a set of workers $D \subseteq W$, and a transfer vector $\check{t}$ such that

$$\begin{aligned} \gamma_{f,D} - \check{t}_{f,D} &\geq \gamma_{f,\mu(f)} - t_{f,\mu(f)}, && \text{and} \\ \alpha_{f,w} + \xi(\check{t}_{f,w}) &\geq \alpha_{\mu(w),w} + \xi(t_{\mu(w),w}) && \forall w \in D, \end{aligned}$$

  with strict inequality for at least one $i \in (\{f\} \cup D)$.

The Kelso and Crawford (1982) (p. 1487) construction of competitive equilibria from strict core allocations then implies that there is some transfer vector $\hat{t}$, having $\hat{t}_{\mu(w),w} = t_{\mu(w),w}$ (for each $w \in W$), such that $[\mu; \hat{t}]$ is stable in our sense.

B.2 Proof of Theorem 1

If $\hat{\mu} = \check{\mu}$, then the theorem is trivially true. Thus, we consider a wage market in which $[\check{\mu}; \check{t}]$ is stable under tax $\check{\tau}$, $[\hat{\mu}; \hat{t}]$ is stable under tax $\hat{\tau}$, $\check{\tau} > \hat{\tau}$, and $\check{\mu} \neq \hat{\mu}$.

The stability conditions for the firms imply that

$$\gamma_{f,\check{\mu}(f)} - \check{t}_{f,\check{\mu}(f)} \geq \gamma_{f,\hat{\mu}(f)} - \check{t}_{f,\hat{\mu}(f)}, \tag{B.1}$$
$$\gamma_{f,\hat{\mu}(f)} - \hat{t}_{f,\hat{\mu}(f)} \geq \gamma_{f,\check{\mu}(f)} - \hat{t}_{f,\check{\mu}(f)}; \tag{B.2}$$

these inequalities together imply that

$$\sum_{f \in F} \left(\check{t}_{f,\hat{\mu}(f)} - \check{t}_{f,\check{\mu}(f)} \right) \geq \sum_{f \in F} \left(\hat{t}_{f,\hat{\mu}(f)} - \hat{t}_{f,\check{\mu}(f)} \right). \tag{B.3}$$

[40] Theorem A.1 of Hatfield et al. (2013) shows that in our setting the Kelso and Crawford (1982) substitutability condition is equivalent to the choice-based substitutability condition of Hatfield and Milgrom (2005) that we describe in the main text: *the availability of new workers cannot make a firm want to hire a worker it would otherwise reject.*

[41] Strictly speaking, Kelso and Crawford (1982) have one technical assumption not present in our framework: they assume that $\gamma_{f,w} + \alpha_{f,w} \geq 0$, in order to ensure that all workers are matched. However, examining the Kelso and Crawford (1982) arguments reveals that this extra assumption is not necessary to ensure that a strict core arrangement exists – the Kelso and Crawford (1982) salary adjustment processes can be started at some arbitrarily low (negative) salary offer and all of the steps and results of Kelso and Crawford (1982) remain valid, with the caveat that some workers may be unmatched at core outcomes.

As the market is a wage market with proportional taxes, we have

$$\xi_\tau(\check{t}_{\check{\mu}(w),w}) = (1-\check{\tau})\check{t}_{\check{\mu}(w),w} \quad \text{and} \quad \xi_\tau(\hat{t}_{\check{\mu}(w),w}) = (1-\hat{\tau})\hat{t}_{\check{\mu}(w),w};$$

hence, the stability conditions for the workers imply that

$$\alpha_{\check{\mu}(w),w} + (1-\check{\tau})\check{t}_{\check{\mu}(w),w} \geq \alpha_{\hat{\mu}(w),w} + (1-\check{\tau})\check{t}_{\hat{\mu}(w),w}, \tag{B.4}$$

$$\alpha_{\hat{\mu}(w),w} + (1-\hat{\tau})\hat{t}_{\hat{\mu}(w),w} \geq \alpha_{\check{\mu}(w),w} + (1-\hat{\tau})\hat{t}_{\check{\mu}(w),w}. \tag{B.5}$$

Summing these inequalities and applying by formula 3.3, we obtain

$$(1-\hat{\tau})\sum_{f\in F}\left(\hat{t}_{f,\hat{\mu}(w)} - \hat{t}_{f,\check{\mu}(f)}\right) \geq (1-\check{\tau})\sum_{f\in F}\left(\check{t}_{f,\hat{\mu}(f)} - \check{t}_{f,\check{\mu}(f)}\right). \tag{B.6}$$

Combining (B.3) and (B.6), we find that

$$(1-\hat{\tau})\sum_{f\in F}\left(\hat{t}_{f,\hat{\mu}(w)} - \hat{t}_{f,\check{\mu}(f)}\right) \geq (1-\check{\tau})\sum_{f\in F}\left(\hat{t}_{f,\hat{\mu}(w)} - \hat{t}_{f,\check{\mu}(f)}\right). \tag{B.7}$$

Since $\hat{\tau} < \check{\tau}$, (B.7) implies that

$$\sum_{f\in F}\left(\hat{t}_{f,\hat{\mu}(f)} - \hat{t}_{f,\check{\mu}(f)}\right) \geq 0. \tag{B.8}$$

Next, using (B.2) and (B.5), we find that

$$\begin{aligned}
\mathfrak{M}(\hat{\mu}) - \mathfrak{M}(\check{\mu}) &= \sum_{f\in F}(\gamma_{f,\hat{\mu}(f)} - \gamma_{f,\check{\mu}(f)}) + \sum_{w\in W}(\alpha_{\hat{\mu}(w),w} - \alpha_{\check{\mu}(w),w}) \\
&\geq \sum_{f\in F}\left(\hat{t}_{f,\hat{\mu}(f)} - \hat{t}_{f,\check{\mu}(f)}\right) - (1-\hat{\tau})\sum_{w\in W}\left(\hat{t}_{\hat{\mu}(w),w} - \hat{t}_{\check{\mu}(w),w}\right), \\
&= \hat{\tau}\sum_{f\in F}\left(\hat{t}_{f,\hat{\mu}(f)} - \hat{t}_{f,\check{\mu}(f)}\right) \geq 0,
\end{aligned}$$

where the final inequality follows from (B.8).

B.3 Proof of Proposition 1

Assume a matching $\check{\mu}$ is stable under tax $\check{\tau}$. In a wage market, if we re-normalize the workers' utilities by dividing by $(1-\check{\tau})$, then any matching that is stable under the renormalized utilities (with no taxation) is also stable under the original utilities and tax $\check{\tau}$: for every worker w and firm f

$$\begin{gathered}
\alpha_{\check{\mu}(w),w} + (1-\check{\tau})\check{t}_{\check{\mu}(w),w} \geq \alpha_{f,w} + (1-\check{\tau})\check{t}_{f,w} \\
\Updownarrow \\
\frac{1}{1-\check{\tau}}\alpha_{\check{\mu}(w),w} + \check{t}_{\hat{\mu}(w),w} \geq \frac{1}{1-\check{\tau}}\alpha_{f,w} + \check{t}_{f,w}.
\end{gathered} \tag{B.9}$$

Combining (B.9) with the firm stability conditions,

$$\gamma_{f,\hat{\mu}(f)} - \hat{t}_{f,\hat{\mu}(f)} \geq \gamma_{f,D} - \hat{t}_{f,D},$$

gives a matching market with quasilinear utility; it is known (see, e.g., Kelso and Crawford (1982); Hatfield et al. (2013)) that in such markets, only an efficient matching can be stable. Thus $\check{\mu}$ must maximize the total of the re-normalized match values,

$$\check{\mu} \in \arg\max_{\{\mu\}} \left[\sum_{f \in F} \gamma_{f,\mu(f)} + \sum_{w \in W} \frac{1}{(1-\check{\tau})} \alpha_{\mu(w),w} \right].$$

For the second and third parts of the proposition, we define a function

$$W(\lambda^\alpha, \lambda^\gamma) \equiv \max_{\{\mu\}} \sum_{f \in F} \left(\lambda^\alpha \alpha_{f,\mu(f)} + \lambda^\gamma \gamma_{f,\mu(f)} \right).$$

which is convex in both λ^α and λ^γ because it is the maximization of a linear function. By the envelope theorem, the derivatives of W are

$$\frac{\partial W(\lambda^\alpha, \lambda^\gamma)}{\partial \lambda^\gamma} = \sum_{f \in F} \gamma_{f,\mu(f)}, \tag{B.10}$$

$$\frac{\partial W(\lambda^\alpha, \lambda^\gamma)}{\partial \lambda^\alpha} = \sum_{f \in F} \alpha_{f,\mu(f)} = \sum_{w \in W} \alpha_{w,\mu(w)}, \tag{B.11}$$

which are the firms' and workers' match values respectively.

It is sufficient to show that, whenever the derivatives (B.10) and (B.11) exist, (a) $\partial W(\lambda^\alpha, \lambda^\gamma)/\partial \lambda^\gamma$ is nondecreasing in λ^γ, and (b) $\partial W(\lambda^\alpha, \lambda^\gamma)/\partial \lambda^\alpha$ is nonincreasing in λ^γ. The first point (a) follows directly from the convexity of W. To see that the cross-derivative is non-positive, (b), we note that W is positive homogeneous of degree 1, so, whenever the derivatives exist, we get by Euler's homogeneous function theorem that

$$W(\lambda^\alpha, \lambda^\gamma) = \lambda^\alpha \frac{\partial W(\lambda^\alpha, \lambda^\gamma)}{\partial \lambda^\alpha} + \lambda^\gamma \frac{\partial W(\lambda^\alpha, \lambda^\gamma)}{\partial \lambda^\gamma}.$$

Hence

$$\begin{aligned} \frac{\partial W(\lambda^\alpha, \lambda^\gamma)}{\partial \lambda^\alpha} &= \frac{1}{\lambda^\alpha} \left(W(\lambda^\alpha, \lambda^\gamma) - \lambda^\gamma \frac{\partial W(\lambda^\alpha, \lambda^\gamma)}{\partial \lambda^\gamma} \right) \\ &= W(1, \lambda^\gamma/\lambda^\alpha) - \frac{\lambda^\gamma}{\lambda^\alpha} \frac{\partial W(1, \lambda^\gamma/\lambda^\alpha)}{\partial \lambda^\gamma}, \end{aligned}$$

which means

$$\frac{\partial}{\partial \lambda^\gamma} \left(\frac{\partial W(\lambda^\alpha, \lambda^\gamma)}{\partial \lambda^\alpha} \right) = \frac{1}{\lambda^\alpha} \frac{\partial W(1, \lambda^\gamma/\lambda^\alpha)}{\partial \lambda^\gamma} - \frac{1}{\lambda^\alpha} \frac{\partial W(1, \lambda^\gamma/\lambda^\alpha)}{\partial \lambda^\gamma} - \frac{\lambda^\gamma}{\lambda^{\alpha^2}} \frac{\partial^2 W(1, \lambda^\gamma/\lambda^\alpha)}{\partial (\lambda^\gamma)^2} < 0,$$

again using the convexity of W. Taking $\lambda^\alpha = 1$ and $\lambda^\gamma = 1 - \check{\tau}$ gives the last part of the proposition.

B.4 Derivation of Equation (A.1)

Summing (B.4) across workers and (B.1) across firms, we find that

$$\sum_{w\in W}\left(\alpha_{\check{\mu}(w),w}-\alpha_{\hat{\mu}(w),w}\right)\geq(1-\check{\tau})\sum_{w\in W}\left(\check{t}_{\hat{\mu}(w),w}-\check{t}_{\check{\mu}(w),w}\right) \tag{B.12}$$

$$\sum_{f\in F}\left(\check{t}_{\hat{\mu}(w),w}-\check{t}_{\check{\mu}(w),w}\right)\geq\sum_{f\in F}\left(\gamma_{f,\hat{\mu}(f)}-\gamma_{f,\check{\mu}(f)}\right). \tag{B.13}$$

As Proposition 1 tells us that

$$\sum_{f\in F}\left(\gamma_{f,\hat{\mu}(f)}-\gamma_{f,\check{\mu}(f)}\right)\geq 0,$$

we can combine (B.12) and (B.13) to get

$$\frac{\sum_{w\in W}\left(\alpha_{\check{\mu}(w),w}-\alpha_{\hat{\mu}(w),w}\right)}{\sum_{f\in F}\left(\gamma_{f,\hat{\mu}(f)}-\gamma_{f,\check{\mu}(f)}\right)}\geq(1-\check{\tau}),$$

(whenever $\sum_{f\in F}\left(\gamma_{f,\hat{\mu}(f)}-\gamma_{f,\check{\mu}(f)}\right)\neq 0$) so that we find

$$\begin{aligned}\check{\tau}&\geq\frac{\sum_{f\in F}\left(\gamma_{f,\hat{\mu}(f)}-\gamma_{f,\check{\mu}(f)}\right)}{\sum_{f\in F}\left(\gamma_{f,\hat{\mu}(f)}-\gamma_{f,\check{\mu}(f)}\right)}+\frac{\sum_{w\in W}\left(\alpha_{\hat{\mu}(w),w}-\alpha_{\check{\mu}(w),w}\right)}{\sum_{f\in F}\left(\gamma_{f,\hat{\mu}(f)}-\gamma_{f,\check{\mu}(f)}\right)}\\&=\frac{\mathfrak{M}(\hat{\mu})-\mathfrak{M}(\check{\mu})}{\sum_{f\in F}\left(\gamma_{f,\hat{\mu}(f)}-\gamma_{f,\check{\mu}(f)}\right)}.\end{aligned} \tag{B.14}$$

B.5 Proof of Proposition 2

First, we show that the arrangements stable under full taxation ($\check{\tau}=1$) cannot Pareto dominate those stable under tax $\hat{\tau}<1$.

Claim 1. *Suppose that $[\hat{\mu};\hat{t}]$ is stable under tax $\hat{\tau}<1$, and that $[\check{\mu};\check{t}]$ is stable under tax $\check{\tau}=1$. Then, $[\check{\mu};\check{t}]$ (under tax $\check{\tau}=1$) cannot Pareto dominate $[\hat{\mu};\hat{t}]$ (under tax $\hat{\tau}<1$).*

Proof. As no transfers get through under full taxation, an arrangement stable under full taxation is most likely to Pareto dominate some other arrangement when all transfers between match partners are 0. Thus, we assume that $\check{t}_{\check{\mu}(w),w}=0$ for each $w\in W$. If $[\check{\mu};\check{t}]$ (under full taxation) Pareto dominates $[\hat{\mu};\hat{t}]$ (under tax $\hat{\tau}$), then

$$\gamma_{f,\check{\mu}(f)}=\gamma_{f,\check{\mu}(f)}-\check{t}_{f,\check{\mu}(f)}\geq\gamma_{f,\hat{\mu}(f)}-\hat{t}_{f,\hat{\mu}(f)}, \tag{B.15}$$

$$\alpha_{\check{\mu}(w),w}=\alpha_{\check{\mu}(w),w}+\xi_{\check{\tau}}(\check{t}_{\check{\mu}(w),w})\geq\alpha_{\hat{\mu}(w),w}+\xi_{\hat{\tau}}(\hat{t}_{\hat{\mu}(w),w}), \tag{B.16}$$

with strict inequality for some f or w. However, stability of $[\hat{\mu};\hat{t}]$ under tax $\hat{\tau}$ implies that

$$\gamma_{f,\hat{\mu}(f)} - \hat{t}_{f,\hat{\mu}(f)} \geq \gamma_{f,\check{\mu}(f)} - \hat{t}_{f,\check{\mu}(f)}, \tag{B.17}$$
$$\alpha_{\hat{\mu}(w),w} + \xi_{\hat{\tau}}(\hat{t}_{\hat{\mu}(w),w}) \geq \alpha_{\check{\mu}(w),w} + \xi_{\hat{\tau}}(\hat{t}_{\check{\mu}(w),w}). \tag{B.18}$$

Combining (B.15) and (B.17) gives

$$0 \geq -\hat{t}_{f,\check{\mu}(f)}, \tag{B.19}$$

for each $f \in F$, while combining (B.16) and (B.18) gives

$$0 \geq \xi_{\hat{\tau}}(\hat{t}_{\check{\mu}(w),w}), \tag{B.20}$$

for each $w \in W$. Strict inequality must hold in (B.19) or (B.20) for some f or w.

In the first of these cases, we have

$$\hat{t}_{f',\check{\mu}(f')} > 0$$

for some $f' \in F$; hence, there exists at least one $w \in \hat{\mu}(f')$ for whom

$$\hat{t}_{\check{\mu}(w),w} > 0. \tag{B.21}$$

But (B.21) contradicts (B.20).

In the second case, we have

$$0 > \xi_{\hat{\tau}}(\hat{t}_{\check{\mu}(w'),w'}), \tag{B.22}$$

for some $w' \in W$. If we take $f = \check{\mu}(w')$, then (B.22) and (B.20) together imply that

$$0 > \sum_{w \in \check{\mu}(f)} \hat{t}_{\check{\mu}(w),w} = \hat{t}_{f,\check{\mu}(f)},$$

contradicting (B.19). □

For $\check{\tau} < 1$, $\xi_{\check{\tau}}(\cdot)$ is strictly increasing and the conclusion of the proposition follows from the following more general result.

Proposition 2'. *Suppose that $\check{\xi}(\cdot)$ is strictly increasing, that $[\hat{\mu};\hat{t}]$ is stable under $\hat{\xi}(\cdot)$, and that $[\check{\mu};\check{t}]$ is stable under $\check{\xi}(\cdot)$, with $\check{\xi}(\cdot) \leq \hat{\xi}(\cdot)$. Then, $[\check{\mu};\check{t}]$ (under $\check{\xi}(\cdot)$) cannot Pareto dominate $[\hat{\mu};\hat{t}]$ (under $\hat{\xi}(\cdot)$).*[42]

[42] We say that an arrangement $[\check{\mu};\check{t}]$ (under $\check{\xi}(\cdot)$) *Pareto dominates* arrangement $[\hat{\mu};\hat{t}]$ under (under $\hat{\xi}(\cdot)$) if

$$\gamma_{f,\check{\mu}(f)} - \check{t}_{f,\check{\mu}(f)} \geq \gamma_{f,\hat{\mu}(f)} - \hat{t}_{f,\hat{\mu}(f)} \qquad \forall f \in F,$$
$$\alpha_{\check{\mu}(w),w} + \check{\xi}(\check{t}_{\check{\mu}(w),w}) \geq \alpha_{\hat{\mu}(w),w} + \hat{\xi}(\hat{t}_{\hat{\mu}(w),w}) \qquad \forall w \in W,$$

with strict inequality for some $i \in F \cup W$.

Proof. Pareto dominance of $[\check{\mu};\check{t}]$ (under $\check{\xi}(\cdot)$) over $[\hat{\mu};\hat{t}]$ (under $\hat{\xi}(\cdot)$) would imply that

$$\gamma_{f,\check{\mu}(f)} - \check{t}_{f,\check{\mu}(f)} \geq \gamma_{f,\hat{\mu}(f)} - \hat{t}_{f,\hat{\mu}(f)}, \tag{B.23}$$

$$\alpha_{\check{\mu}(w),w} + \check{\xi}(\check{t}_{\check{\mu}(w),w}) \geq \alpha_{\hat{\mu}(w),w} + \hat{\xi}(\hat{t}_{\hat{\mu}(w),w}), \tag{B.24}$$

with strict inequality for some f or w. However, stability of $[\hat{\mu};\hat{t}]$ under $\hat{\xi}(\cdot)$ implies that

$$\gamma_{f,\hat{\mu}(f)} - \hat{t}_{f,\hat{\mu}(f)} \geq \gamma_{f,\check{\mu}(f)} - \hat{t}_{f,\check{\mu}(f)}, \tag{B.25}$$

$$\alpha_{\hat{\mu}(w),w} + \hat{\xi}(\hat{t}_{\hat{\mu}(w),w}) \geq \alpha_{\check{\mu}(w),w} + \hat{\xi}(\hat{t}_{\check{\mu}(w),w}) \geq \alpha_{\check{\mu}(w),w} + \check{\xi}(\hat{t}_{\check{\mu}(w),w}), \tag{B.26}$$

where the second inequality in (B.26) follows from the fact that $\hat{\xi}(\cdot) \geq \check{\xi}(\cdot)$.

Combining (B.23) and (B.25) gives

$$\hat{t}_{f,\check{\mu}(f)} \geq \check{t}_{f,\check{\mu}(f)}, \tag{B.27}$$

for each $f \in F$, while combining (B.24) and (B.26) gives

$$\begin{aligned} \check{\xi}(\check{t}_{\check{\mu}(w),w}) &\geq \check{\xi}(\hat{t}_{\check{\mu}(w),w}) \\ \check{t}_{\check{\mu}(w),w} &\geq \hat{t}_{\check{\mu}(w),w} \end{aligned} \tag{B.28}$$

for each $w \in W$, where the second line of (B.28) follows from the fact that $\check{\xi}(\cdot)$ is strictly increasing. Strict inequality must hold in (B.27) or (B.28) for some f or w.

In the first of these cases, we have

$$\hat{t}_{f',\check{\mu}(f')} > \check{t}_{f',\check{\mu}(f')}$$

for some $f' \in F$; hence, there exists at least one $w \in \hat{\mu}(f')$ for whom

$$\hat{t}_{\check{\mu}(w),w} > \check{t}_{\check{\mu}(w),w}. \tag{B.29}$$

But (B.29) contradicts (B.28).

In the second case, we have

$$\check{t}_{\check{\mu}(w'),w'} > \hat{t}_{\check{\mu}(w'),w'} \tag{B.30}$$

for some $w' \in W$. If we take $f = \check{\mu}(w')$, then (B.30) and (B.28) together imply that

$$\sum_{w \in \check{\mu}(f)} \check{t}_{\check{\mu}(w),w} > \sum_{w \in \check{\mu}(f)} \hat{t}_{\check{\mu}(w),w};$$

hence, we find that

$$\check{t}_{f,\check{\mu}(f)} > \hat{t}_{f,\check{\mu}(f)},$$

contradicting (B.27). $\square$

B.6 Proofs of Proposition 4 and Corollary 1

Suppose that in a wage market, both $[\check{\mu};\check{t}]$ and $[\hat{\mu};\hat{t}]$ are stable under tax τ. The stability conditions for the firms imply that

$$\gamma_{f,\check{\mu}(f)} - \check{t}_{f,\check{\mu}(f)} \geq \gamma_{f,\hat{\mu}(f)} - \check{t}_{f,\hat{\mu}(f)}, \tag{B.31}$$

$$\gamma_{f,\check{\mu}(f)} - \hat{t}_{f,\check{\mu}(f)} \leq \gamma_{f,\hat{\mu}(f)} - \hat{t}_{f,\hat{\mu}(f)}, \tag{B.32}$$

so that

$$\check{t}_{f,\hat{\mu}(f)} - \check{t}_{f,\check{\mu}(f)} \geq \hat{t}_{f,\hat{\mu}(f)} - \hat{t}_{f,\check{\mu}(f)}. \tag{B.33}$$

Meanwhile, the stability conditions for the workers imply that

$$\alpha_{\check{\mu}(w),w} + (1-\tau)\check{t}_{\check{\mu}(w),w} \geq \alpha_{\hat{\mu}(w),w} + (1-\tau)\check{t}_{\hat{\mu}(w),w}, \tag{B.34}$$

$$\alpha_{\check{\mu}(w),w} + (1-\tau)\hat{t}_{\check{\mu}(w),w} \leq \alpha_{\hat{\mu}(w),w} + (1-\tau)\hat{t}_{\hat{\mu}(w),w}, \tag{B.35}$$

so that

$$(1-\tau)(\check{t}_{\hat{\mu}(w),w} - \check{t}_{\check{\mu}(w),w}) \leq (1-\tau)(\hat{t}_{\hat{\mu}(w),w} - \hat{t}_{\check{\mu}(w),w}). \tag{B.36}$$

Summing (B.33) and (B.36) across agents and using formula 3.3, we find that

$$\sum_{f\in F}\left(\check{t}_{f,\hat{\mu}(f)} - \check{t}_{f,\check{\mu}(f)}\right) = \sum_{f\in F}\left(\hat{t}_{f,\hat{\mu}(f)} - \hat{t}_{f,\check{\mu}(f)}\right). \tag{B.37}$$

For this equality to hold, we must have equality in (B.33) for each $f \in F$, implying equality in (B.31) and (B.32), for each $f \in F$. Similarly, (B.37) requires that (B.36) hold with equality for each $w \in W$, which implies equality in (B.34) and (B.35), for each $w \in W$. Combining these equalities, and summing across workers $w \in W$, shows that

$$\begin{aligned}\sum_{w\in W}\left(\alpha_{\check{\mu}(w),w} - \alpha_{\hat{\mu}(w),w}\right) &= (1-\tau)\sum_{f\in F}\left(\check{t}_{f,\hat{\mu}(f)} - \check{t}_{f,\check{\mu}(f)}\right),\\ &= (1-\tau)\sum_{f\in F}\left(\gamma_{f,\hat{\mu}(f)} - \gamma_{f,\check{\mu}(f)}\right).\end{aligned} \tag{B.38}$$

If the firms are not indifferent in aggregate between $\check{\mu}$ and $\hat{\mu}$, so that

$$\sum_{f\in F}\left(\gamma_{f,\hat{\mu}(f)} - \gamma_{f,\check{\mu}(f)}\right) \neq 0, \tag{B.39}$$

we have,

$$\tau = 1 + \frac{\sum_{w\in W}\left(\alpha_{\check{\mu}(w),w} - \alpha_{\hat{\mu}(w),w}\right)}{\sum_{f\in F}\left(\gamma_{f,\check{\mu}(f)} - \gamma_{f,\hat{\mu}(f)}\right)}, \tag{B.40}$$

thereby showing Proposition 4.

To see Corollary 1, observe that (B.40) pins down a unique tax rate in the case that (B.39) holds. Thus, if there are two tax rates under which matchings $\check{\mu}$ and $\hat{\mu}$ are both stable, then we must have

$$\sum_{f \in F} \left(\gamma_{f,\hat{\mu}(f)} - \gamma_{f,\check{\mu}(f)} \right) = 0. \tag{B.41}$$

But then, we also have

$$\sum_{w \in W} \left(\alpha_{\hat{\mu}(w),w} - \alpha_{\check{\mu}(w),w} \right) = 0, \tag{B.42}$$

by (B.38). Together (B.41) and (B.42) imply $\mathfrak{M}(\hat{\mu}) - \mathfrak{M}(\check{\mu}) = 0$, as desired.

C Simulations of Non-Wage Markets

We run 500 simulations of a one-to-one market with twenty agents on each side of the market and match values independently and identically distributed according to a uniform distribution on $[-.5, .5]$. We vary the tax rate, τ, from 0 to .99 in increments of .01. For each tax rate, we find the manager-optimal stable arrangement and calculate the total match value – if there are multiple stable arrangements, the manager-optimal arrangement is the one preferred by all managers. Non-monotonicities in the total match value of stable matchings appear in over half of the markets (55%). There may be additional non-monotonicities that we do not observe because we cannot vary τ continuously. However, the non-monotonicities we fail to observe necessarily occur over very small ranges of τ, as we observe all non-monotonicities that persist over values of τ for a range of .01 or more.

Most markets have relatively small losses from non-monotonicity, mostly occurring at high tax rates, but some have dramatic non-monotonicities. Table 3 summarizes the non-monotonocities arising in our simulations. Row 1 shows the fraction of markets that have non-monotonicities in a given tax rate range. While the majority of non-monotonicities occur at very high tax rates, 10% of our simulation markets have non-monotonicities at tax rates below 50%. Row 2 gives the (normalized) average size of the non-monotonicities in each tax rate range. Again, we see that non-monotonicities are most significant for high tax rates. Row 3 incorporates information on the persistence of non-monotonicities by computing the fraction of the deadweight loss from taxation that is due to a non-monotonicity. The fraction his is relatively high for lower tax rates because there is less total deadweight loss at those tax rates.

Overall, our simulations suggest non-monotonicities in the tax rate are not just artifacts of example selection. However, they also suggest that non-monotonicities are relatively rare at more realistic tax rates ($\tau \in [0, .5)$) and tend not to persist over large ranges of τ.[43]

[43] Increasing the sample size does not appear to decrease the frequency or importance of non-monotonicities.

Table 3: Summary of the non-monotonicities arising in simulated markets.

	Range of τ				
	$[0,.25)$	$[.25,.5)$	$[.5,.75)$	$[.75,1)$	All τ
Fraction of markets with non-monotonicity	0.006	0.088	0.190	0.394	0.548
Avg size of non-monotonicity, as fraction of range	0.021	0.066	0.111	0.140	0.120
Fraction of deadweight loss from taxation due to non-monotonicity	0.076	0.070	0.051	0.027	0.037

Note: The table summarizes 500 simulations of one-to-one matching markets with 20 agents on each side of the market. All agents' match values are independently and identically distributed according to a uniform distribution on $[-.5,.5]$. We vary the tax rate, τ, from 0 to .99 in increments of .01. For each tax rate, we find the manager-optimal stable arrangement and calculate the total match value. Row 1 presents the fraction of markets that have non-monotonicities in a given tax rate range. Row 2 presents the average size of non-monotonicities within each range, normalized as a fraction of the (within-market) gap between the highest and lowest total stable match values calculated for any tax rate. Row 3 presents the average fraction of taxation deadweight loss that is due to non-monotonicity, across all markets. The deadweight loss from non-monotonicity is computed for each tax rate τ as the difference between the highest total match value for a tax rate $\tilde{\tau} \geq \tau$ and the total match value under tax rate τ; it is divided by the total deadweight loss from taxation at tax rate τ, which is computed as the difference in total match value between the efficient matching and the matching stable under tax rate τ.

D Lump Sum Taxation

While not typically phrased in the exact language of taxation, lump sum taxes are prevalent in labor markets. They can take the form of costs for hiring (e.g., employee health care costs) or for entering employment (e.g., licensing requirements). In the marriage market context, lump sum taxes can take the form of marriage license fees or tax penalties for marriage.

D.1 Lump Sum Taxation of Transfers

We first consider a lump sum tax that is levied only on (nonzero) transfers between match partners.[44] Such a *lump sum tax on transfers*, ℓ, corresponds to the transfer function

$$\xi_\ell^{\mathsf{lump}}(t_{f,w}) \equiv \begin{cases} t_{f,w} - \ell & t_{f,w} \neq 0, \\ t_{f,w} & t_{f,w} = 0. \end{cases}$$

shown in Figure 9. Under this tax structure, the case $\ell = 0$ corresponds to the standard (Shapley and Shubik (1971)) model of matching with transfers and the case $\ell = \infty$ corre-

[44] An alternative approach to lump sum taxation, which we discuss in the next section, imposes a flat fee on all matches.

sponds to (Gale and Shapley (1962)) matching without transfers.

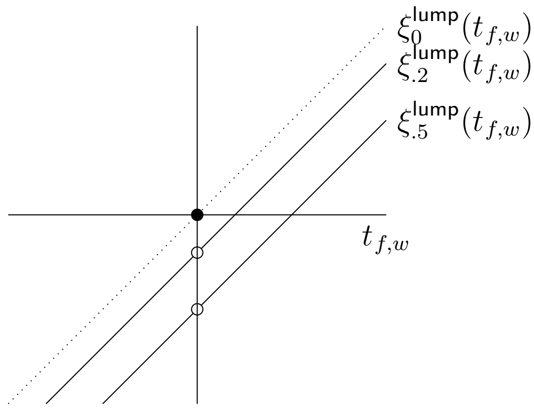


Figure 9: Transfer function $\xi_\ell^{\mathsf{lump}}(\cdot)$.

We say that an arrangement or matching is *stable under lump sum tax* ℓ if it is stable given transfer function $\xi_\ell^{\mathsf{lump}}(\cdot)$.

A lump sum tax on transfers has an extensive margin effect that makes being unmatched more attractive relative to matching with a transfer. In non-wage markets,[45] a lump sum tax on transfers can also encourage matchings in which transfers are unnecessary.[46] As our next example illustrates, this second distortion can cause the total match value of stable matchings to be non-monotonic in the size of the lump sum tax.

Example 3 (Non-monotonicity). *Consider a one-to-one market with two firms,* $F = \{f_1, f_2\}$, *two workers,* $W = \{w_1, w_2\}$, *and match values as pictured in Figure 10a. Worker* w_1 *likes* f_1 *– who has a strong preference for* w_2 *– but* w_2 *prefers* f_2*. When transfers are not allowed (or when there is a high lump sum tax on transfers,* $\ell \geq 18$*), the only stable matching is the matching* μ_1 *in which* $\mu_1(f_1) = w_1$ *and* $\mu_1(f_2) = w_2$*, as shown in Figure 10b. This matching yields total match value of* $\mathfrak{M}(\mu_1) = 22$.

When the lump sum tax is lowered to $\ell = 12$*, only the matching* μ_2 *is stable, where* $\mu_2(f_1) = w_2$ *and* w_1 *and* f_2 *are unmatched; this matching gives a total match value* $\mathfrak{M}(\mu_2) = 19$*, as shown in Figure 10d. When* $\ell = 12$*, the tax is low enough that* f_1 *can convince* w_2 *to match with him, but not low enough for* w_1 *to hold onto* f_1 *when it has the option of matching with* w_2 *(or* f_2 *to hold onto* w_2*). Lowering the lump sum tax from* 20 *to* 12 *decreases the total match value of the stable matching and decreases the number of agents matched.*

We use simulations to confirm that Example 3 is not an exceptional case. In the same small markets described in Appendix C, with utilities uniformly distribution on $[-.5, .5]$,

[45] Since it is difficult to observe transfers in non-wage markets, such as marriage markets, it is somewhat hard to imagine taxing them. Nevertheless, lump-sum taxes on transfers could correspond to instituting a lump sum tax on gifts between spouses, and flat fees for matching could correspond to requiring marriage license fees.

[46] Consider the case of balanced one-to-one matching markets. In such markets, lump sum taxes on transfers promotes pairing (f, w) in which the match value $\gamma_{f,w} + \alpha_{f,w}$ is evenly distributed between the two partners ($\gamma_{f,w} \approx \alpha_{f,w}$), so that transfers are unnecessary.

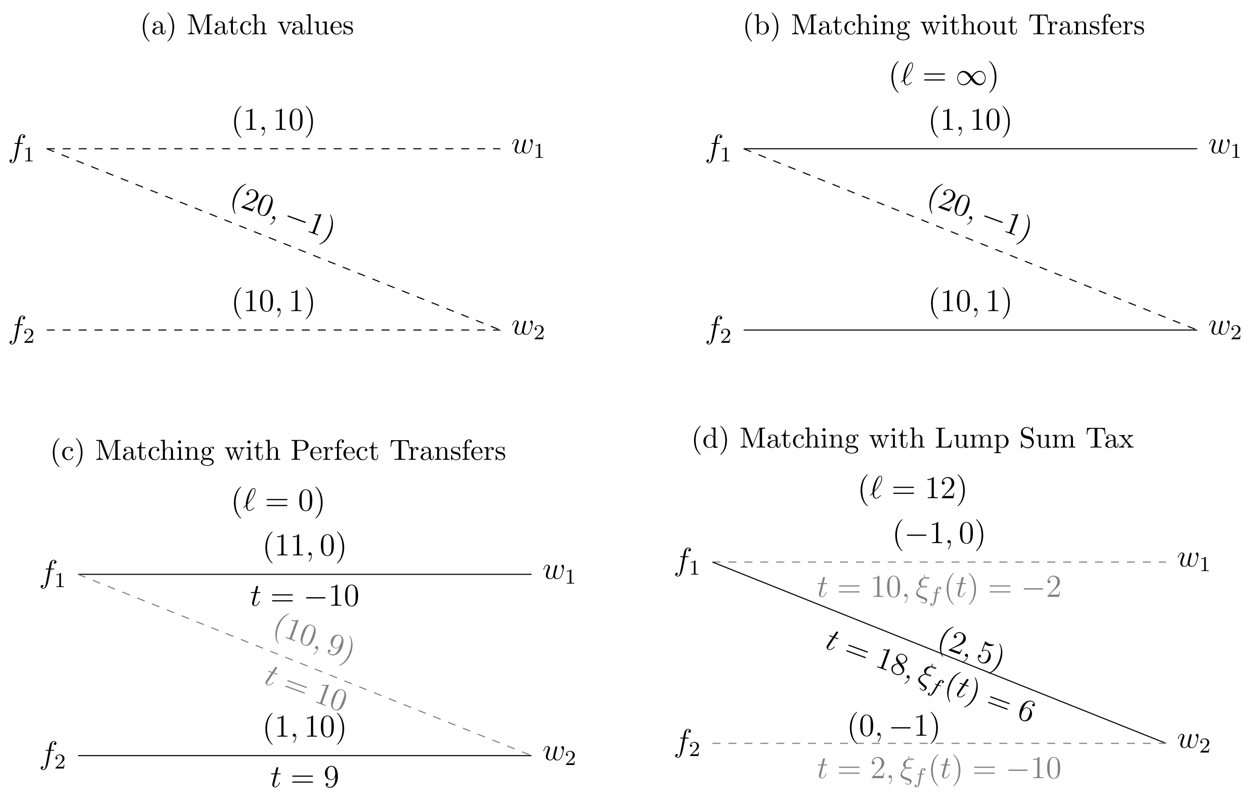


Figure 10: Example 3 – Non-monotonicity under a lump sum tax on transfers.

Note: Utilities, net of transfers, are above the lines (firm's, worker's). Possible supporting transfers (when applicable) are below the lines. Solid lines indicate the stable matching.

we find that match value is non-monotonic in the lump sum tax in 61% of our simulated markets.

In a *strictly positive wage market* worker match values are strictly negative (instead of just non-positive). In strictly positive wage markets, all matchings require a transfer, so a lump sum tax on transfers does not distort agents' preferences among match partners – for a given transfer vector, if a worker prefers firm f_1 to f_2 without a tax, then that worker also prefers f_1 to f_2 under a lump sum tax. Thus, in strictly positive wage markets, the matching distortion of the lump sum tax is only on the extensive margin – the decision of *whether* to match – under a higher lump sum tax, fewer agents find matching desirable. This intuition that lump sum taxes work on the extensive margin is captured in the following lemma, where we use $\#(\mu)$ to denote the number of workers matched in matching μ

Lemma 1. *In strictly positive wage markets, reduction in a lump sum tax on transfers (weakly) increases the number of workers matched in stable matchings. That is, if matching $\check{\mu}$ is stable under lump sum tax $\check{\ell}$, matching $\hat{\mu}$ is stable under lump sum tax $\hat{\ell}$, and $\hat{\ell} < \check{\ell}$, then*

$$\#(\hat{\mu}) \geq \#(\check{\mu}).$$

In non-wage markets, the conclusion of Lemma 1 is not true, in general, because distortion among match partners can dominate the extensive margin effect, as in Example 3.

As lump sum taxes do not distort among match partners in strictly positive wage markets, they can only reduce the efficiency of stable matchings in such markets by reducing the number of workers matched. This idea that the distortion must be on the extensive margin, combined with Lemma 1, gives the following result.

Theorem 2. *In strictly positive wage markets, a reduction in a lump sum tax on transfers (weakly) increases the total match value of stable matchings. That is, if $\check{\mu}$ is stable under lump sum tax $\check{\ell}$, $\hat{\mu}$ is stable under lump sum tax $\hat{\ell}$, and $\hat{\ell} < \check{\ell}$, then*

$$\mathfrak{M}(\hat{\mu}) \geq \mathfrak{M}(\check{\mu}).$$

Theorem 2 indicates that in strictly positive wage markets, match value increases monotonically as lump sum taxation decreases.

In strictly positive wage markets, we can also bound the total match value loss from a given lump sum tax.

Proposition 5. *In a strictly positive wage market, let $\hat{\mu}$ be an efficient matching, and let $\check{\mu}$ be stable under lump sum tax on transfers $\check{\ell}$. Then,*

$$0 \leq \mathfrak{M}(\hat{\mu}) - \mathfrak{M}(\check{\mu}) \leq \check{\ell} \cdot (\#(\hat{\mu}) - \#(\check{\mu})).$$

The intuition for Proposition 5 is that since the workers unmatched under a lump sum tax of $\check{\ell}$ have negative surplus from matching under that lump sum tax, their surplus from matching could not be more than $\check{\ell}$. So the change in total value is less than the change in the number of unmatched workers times a maximum surplus of $\check{\ell}$ per worker.

Finally, we can show that, for a *fixed* limit on the number of workers matched in the presence of a lump sum tax, stable matchings in strictly positive wage markets must generate the maximal match value possible.

Proposition 6. *In a strictly positive wage market, a matching $\breve{\mu}$ can be stable under a lump sum tax on transfers only if*

$$\breve{\mu} \in \underset{\{\mu : \#(\mu) \leq \#(\breve{\mu})\}}{\arg\max} \left[\mathfrak{M}(\mu)\right].$$

Proposition 6 shows that a lump sum tax is an efficient way for a market designer to limit the number of matches (in strictly positive wage markets): the matchings stable under lump sum taxation have maximal value, given the tax's implied limit on the number of agents matched. Analogously, if a market designer wants to encourage matches, a lump-sum subsidy will maximize total match value for a given (subsidy-induced) lower bound on the number of agents matched. For example, Proposition 6 suggests that if a government wants to use tuition subsidies to encourage people to go to school, then uniform tuition subsidies are more efficient than subsidies proportional to the cost of tuition.

D.2 Lump Sum Taxation of Matches

Some fee structures tax *all* pairings, rather than just those that include nonzero transfers. Such *flat fees for matching* can also be interpreted in the language of taxation: they correspond to the transfer function

$$\xi_{\ell}^{\mathsf{fee}}(t_{f,w}) \equiv t_{f,w} - \ell.$$

Figure 11 shows this transfer function for different levels of ℓ.

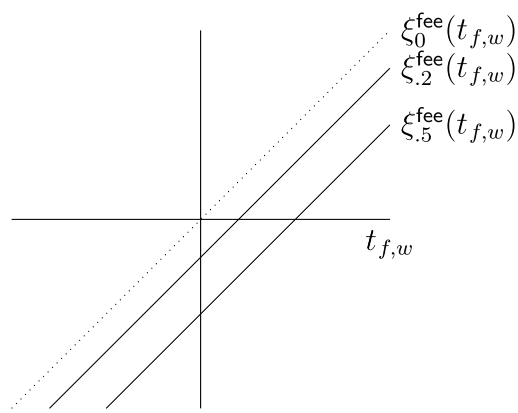


Figure 11: Transfer function $\xi_{\ell}^{\mathsf{fee}}(\cdot)$.

Unlike lump sum taxes on transfers, flat fees for matching never distort among match partners – even in non-wage markets. Flat fees for matching only have extensive margin effects, and thus markets with such fees are similar to strictly positive wage markets with lump sum taxes on transfers.[47] As shown below, the conclusions of Lemma 1, Theorem 2, and Propositions 5 and 6 *always* hold in markets with flat fees for matching.

[47]Indeed, in strictly positive wage markets, lump sum taxation of transfers is equivalent to lump sum taxation of matchings because workers never match without receiving a strictly positive transfer.

D.3 Proofs

Proof of Lemma 1

In a strictly positive wage market, all matches are accompanied by a strictly positive transfer; hence, a lump sum tax on transfers is equivalent to a flat fee for matching. Thus, Lemma 1 follows from the following slightly more general result.

Here and hereafter, we say that an arrangement or matching is *stable under flat fee ℓ* if it is stable given transfer function $\xi_\ell^{\mathsf{fee}}(\cdot)$.

Lemma 1'. *Reduction of a flat fee for matching (weakly) increases the number of workers matched in stable matchings. That is, if matching $\check{\mu}$ is stable under flat fee $\check{\ell}$, matching $\hat{\mu}$ is stable under flat fee $\hat{\ell}$, and $\hat{\ell} < \check{\ell}$, then*

$$\#(\hat{\mu}) \geq \#(\check{\mu}),$$

where $\#(\mu)$ denotes the number of workers matched in matching μ.

Proof. As $[\check{\mu}; \check{t}]$ is stable under flat fee $\check{\ell}$, we have

$$\gamma_{f,\check{\mu}(f)} - \check{t}_{f,\check{\mu}(f)} \geq \gamma_{f,\hat{\mu}(f)} - \check{t}_{f,\hat{\mu}(f)}$$
$$\alpha_{\check{\mu}(w),w} + \check{t}_{\check{\mu}(w),w} - \check{\ell} \cdot \{1_{\check{\mu}(w)\neq w}\} \geq \alpha_{\hat{\mu}(w),w} + \check{t}_{\hat{\mu}(w),w} - \check{\ell} \cdot \{1_{\hat{\mu}(w)\neq w}\}$$

where $\{1_{\mu(w)\neq w}\}$ is an indicator function that equals 1 if w is matched in matching μ and 0 if w is unmatched in matching μ. Summing these inequalities across agents, and formula 3.3, we find that

$$\sum_{f\in F} \left(\gamma_{f,\check{\mu}(f)} - \gamma_{f,\hat{\mu}(f)}\right) + \sum_{w\in W} \left(\alpha_{\check{\mu}(w),w} - \alpha_{\hat{\mu}(w),w}\right) + \check{\ell} \cdot (\#(\hat{\mu}) - \#(\check{\mu})) \geq 0. \tag{D.1}$$

Similarly, as $[\hat{\mu}; \hat{t}]$ is stable under flat fee $\hat{\ell}$,

$$\gamma_{f,\hat{\mu}(f)} - \hat{t}_{f,\hat{\mu}(f)} \geq \gamma_{f,\check{\mu}(f)} - \hat{t}_{f,\check{\mu}(f)}$$
$$\alpha_{\hat{\mu}(w),w} + \hat{t}_{\hat{\mu}(w),w} - \hat{\ell} \cdot \{1_{\hat{\mu}(w)\neq w}\} \geq \alpha_{\check{\mu}(w),w} + \hat{t}_{\check{\mu}(w),w} - \hat{\ell} \cdot \{1_{\check{\mu}(w)\neq w}\};$$

these inequalities yield

$$\sum_{f\in F} \left(\gamma_{f,\hat{\mu}(f)} - \gamma_{f,\check{\mu}(f)}\right) + \sum_{w\in W} \left(\alpha_{\hat{\mu}(w),w} - \alpha_{\check{\mu}(w),w}\right) + \hat{\ell} \cdot (\#(\check{\mu}) - \#(\hat{\mu})) \geq 0. \tag{D.2}$$

upon summation.

Adding (D.1) and (D.2) shows that

$$(\check{\ell} - \hat{\ell})(\#(\hat{\mu}) - \#(\check{\mu})) \geq 0.$$

Thus, if $\check{\ell} > \hat{\ell}$, we must have $\#(\hat{\mu}) \geq \#(\check{\mu})$; this proves the result. $\square$

Proof of Theorem 2

As in the proof of Lemma 1, Theorem 2 follows from the following slightly more general result.

Theorem 2'. *A reduction in a flat fee for matching (weakly) increases the total match value of stable matchings. That is, if $\check{\mu}$ is stable under flat fee $\check{\ell}$, $\hat{\mu}$ is stable under flat fee $\hat{\ell}$, and $\hat{\ell} < \check{\ell}$, then*

$$\mathfrak{M}(\hat{\mu}) \geq \mathfrak{M}(\check{\mu}).$$

Proof. Using (D.2) and Lemma 1', we find that

$$\mathfrak{M}(\hat{\mu}) - \mathfrak{M}(\check{\mu}) = \sum_{f \in F} \left(\gamma_{f,\hat{\mu}(f)} - \gamma_{f,\check{\mu}(f)}\right) + \sum_{w \in W} \left(\alpha_{\hat{\mu}(w),w} - \alpha_{\check{\mu}(w),w}\right) \geq \hat{\ell} \cdot (\#(\hat{\mu}) - \#(\check{\mu})) \geq 0;$$

this proves Theorem 2'. □

Proof of Proposition 5

As in the proof of Lemma 1, Proposition 5 follows from the following slightly more general result.

Proposition 5'. *Let $\hat{\mu}$ be an efficient matching, and let $\check{\mu}$ be stable under flat fee $\check{\ell}$. Then,*

$$0 \leq \mathfrak{M}(\hat{\mu}) - \mathfrak{M}(\check{\mu}) \leq \check{\ell} \cdot (\#(\hat{\mu}) - \#(\check{\mu})).$$

Proof. The Proposition is immediate from (D.1). □

Proof of Proposition 6

As in the proof of Lemma 1, Proposition 6 follows from the following slightly more general result.

Proposition 6'. *A matching $\check{\mu}$ can be stable under a flat fee only if*

$$\check{\mu} \in \underset{\{\mu \,:\, \#(\mu) \leq \#(\check{\mu})\}}{\arg\max} \left\{\mathfrak{M}(\mu)\right\}.$$

Proof. From (D.1), we see that if $[\check{\mu}; \check{t}]$ is stable under flat fee $\check{\ell}$, then for any matching $\hat{\mu} \neq \check{\mu}$,

$$\mathfrak{M}(\check{\mu}) - \mathfrak{M}(\hat{\mu}) + \check{\ell} \cdot (\#(\hat{\mu}) - \#(\check{\mu})) \geq 0. \tag{D.3}$$

If fewer workers are matched in $\hat{\mu}$ than in $\check{\mu}$ (i.e. $\#(\check{\mu}) \geq \#(\hat{\mu})$), (D.3) implies that

$$\mathfrak{M}(\check{\mu}) - \mathfrak{M}(\hat{\mu}) \geq \hat{\ell} \cdot (\#(\check{\mu}) - \#(\hat{\mu})) \geq 0,$$

so that $\check{\mu}$ must have higher total match value than $\hat{\mu}$. □

References

247Sports, "2013 football recruiting team rankings," 2013. Online; accessed 29-July-2016.

Arcidiacono, Peter, Andrew Beauchamp, and Marjorie McElroy, "Terms of endearment: An equilibrium model of sex and matching," 2011. NBER Working Paper No. 16517.

Becker, Gary S., "A theory of marriage: Part II," *Journal of Political Economy*, 1974, *82*, 11–26.

Blundell, Richard, Alan Duncan, and Costas Meghir, "Estimating labor supply responses using tax reforms," *Econometrica*, 1998, *66*, 827–861.

Boadway, R., M. Marchand, and P. Pestieau, "Optimal linear income taxation in models with occupational choice," *Journal of Public Economics*, 1991, *46* (2), 133–162.

Chetty, Raj, "Is the taxable income elasticity sufficient to calculate deadweight loss? The implications of evasion and avoidance," *American Economic Journal: Economic Policy*, 2009, pp. 31–52.

Chiappori, Pierre-André and Bernard Salanié, "The econometrics of matching models," *Journal of Economic Literature*, 2016, *54* (3), 832–61.

Choo, Eugene and Aloysius Siow, "Who marries whom and why," *Journal of Political Economy*, 2006, *114* (1), 175–201.

Crawford, Vincent P. and Elsie Marie Knoer, "Job matching with heterogeneous firms and workers," *Econometrica*, 1981, *49*, 437–450.

Dupuy, Arnaud and Alfred Galichon, "A note on the estimation of job amenities and labor productivity," 2017. IZA working paper 10909.

Epstein, Brendan and Ryan Nunn, "Taxation, match quality and social welfare," 2013. Federal Reserve System International Finance Discussion Paper No. 1079.

ESPN, "ESPN football power index-2013," 2013. Online; accessed 29-July-2016.

Feldstein, Martin, "Tax avoidance and the deadweight loss of the income tax," *The Review of Economics and Statistics*, 1999, *81* (4), 674–680.

Fok, Dennis, Richard Paap, and Bram Van DijkFok, "A Rank-ordered Logit Model with Unobserved Heterogeneity in Ranking Capabilities," *Journal of Applied Econometrics*, 2012, *27* (5), 831–846.

Fréchette, Guillaume R, Alvin E. Roth, and M. Utku Ünver, "Unraveling yields inefficient matchings: Evidence from post-season college football bowls," *RAND Journal of Economics*, 2007, *38* (4), 967–982.

Gale, David and Lloyd S. Shapley, "College admissions and the stability of marriage," *American Mathematical Monthly*, 1962, *69*, 9–15.

Galichon, Alfred and Bernard Salanié, "Cupid's invisible hand: Social surplus and identification in matching models," 2014. Mimeo.

_ , Scott Kominers, and Simon Weber, "Costly concessions: An empirical framework for matching with imperfectly transferable utility," 2017. Mimeo.

Gentry, William M. and R. Glenn Hubbard, "The effects of progressive income taxation on job turnover," *Journal of Public Economics*, 2004, *88* (11), 2301–2322.

Hatfield, John William and Paul Milgrom, "Matching with contracts," *American Economic Review*, 2005, *95*, 913–935.

— , **Scott Duke Kominers, Alexandru Nichifor, Michael Ostrovsky, and Alexander Westkamp**, "Stability and competitive equilibrium in trading networks," *Journal of Political Economy*, 2013, *121*, 966–1005.

Holzner, Christian and Andrey Launov, "Do employed workers search efficiently? Implications for the income tax system," 2012. Mimeo.

Kelso, Alexander S. and Vincent P. Crawford, "Job matching, coalition formation, and gross substitutes," *Econometrica*, 1982, *50*, 1483–1504.

Koopmans, Tjalling C. and Martin Beckmann, "Assignment problems and the location of economic activities," *Econometrica*, 1957, *25* (1), 53–76.

Legros, Patrick and Andrew F. Newman, "Beauty is a beast, frog is a prince: Assortative matching with nontransferabilities," *Econometrica*, 2007, *75*, 1073–1102.

Lockwood, Benjamin B., Charles G. Nathanson, and E. Glen Weyl, "Taxation and the allocation of talent," *Journal of Political Economy*, 2017, *125* (5), 1635–1682.

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green, *Microeconomic Theory*, Oxford University Press, 1995.

McFadden, Daniel, "Quantitative Methods for Analyzing Travel Behaviour of Individuals: Some Recent Developments," Technical Report 1977.

Meyer, Bruce D., "Labor supply at the extensive and intensive margins: The EITC, welfare, and hours worked," *American Economic Review*, 2002, *92* (2), 373–379.

National Collegiate Athletics Association, "Winning percentage," 2013. Online; accessed 29-July-2016.

Office of Postsecondary Education, 2013. https://ope.ed.gov/athletics; accessed 29-July-2016.

Parker, Simon C., "Does tax evasion affect occupational choice?," *Oxford Bulletin of Economics and Statistics*, 2003, *65* (3), 379–394.

Powell, David and Hui Shan, "Income taxes, compensating differentials, and occupational choice: How taxes distort the wage-amenity decision," *American Economic Journal: Economic Policy*, 2012, *4* (1), 224–247.

Quinzii, Martine, "Core and competitive equilibria with indivisibilities," *International Journal of Game Theory*, 1984, *135*, 41–60.

Roth, Alvin E., "The economics of matching: Stability and incentives," *Mathematics of Operations Research*, 1982, *7*, 617–628.

— **and Marilda A. O. Sotomayor**, "The college admissions problem revisited," *Econometrica*, 1989, *57* (3), 559–570.

Rothschild, Casey and Florian Scheuer, "Redistributive taxation in the Roy model," *The Quarterly Journal of Economics*, 2012, pp. 623–668.

Saez, Emmanuel, "Optimal income transfer programs: intensive versus extensive labor supply responses," *Quarterly Journal of Economics*, 2002, *117* (3), 1039–1073.

— , "Reported incomes and marginal tax rates, 1960-2000: Evidence and policy implications," in James Poterba, ed., *Tax Policy and the Economy*, MIT Press, 2004.

Scheuer, Florian and Iván Werning, "The taxation of superstars," *Quarterly Journal of Economics*, 2017, *132* (1), 211–270.

Shapley, Lloyd S. and Martin Shubik, "The assignment game I: The core," *International Journal of Game Theory*, 1971, *1*, 111–130.

Sheshinski, Eytan, "Note on income taxation and occupational choice," 2003. CESifo Working Paper No. 880.

Tax Foundation, "Federal and state level tax information 2013," 2013. Online; accessed 17-November-2016.

USA Today, "Coaches salaries of the division I FBS," 2013. Online; accessed 29-July-2016.

Wikipedia, "Head coaches past performance at start of 2013 season," 2013. Online; accessed 01-June-2015.