Investment in matching markets*

Matthew Elliott† Eduard Talamàs‡

1 Introduction

Non-contractible investments are often made before entering a matching market. Firms create vacancies before hiring workers, while workers acquire human capital before looking for a job; manufacturers and suppliers form specific relationships with each other; entrepreneurs invest in their businesses before securing support from venture capitalists; and so on.

The hold-up problem can be a significant barrier to efficient investments in these settings. Investments are often sunk by the time people bargain over how to share the returns of these investments. Since bargaining need not give each person the full surplus generated by her investment, private and social investment incentives need not be aligned.

In this chapter, we analyze investments prior to participating in a matching market—paying particular attention to the hold-up problem. We focus on finite one-to-one matching markets with transfers. Our objective is to highlight some of the key forces that shape investment decisions in matching markets—rather than to survey the literature. The outline is as follows.

Section 2 illustrates the main questions that we focus on in the context of a market that matches designers with design studios. Section 3 describes our benchmark model, which treats investment decisions as a noncooperative game and the matching/bargaining process as a cooperative game. Section 4 discusses the conditions that need to hold to align private and social investment incentives for a given person.

Building on the analysis in Section 4, Section 5 presents and discusses the main result of this chapter: Everyone on one side of the market simultaneously obtains her full marginal product—and hence has good investment incentives—if the bargaining process in the matching stage ends up selecting this side's optimal core outcome. Section 6 provides a new proof of this classic result.

* *Date printed*: May 20, 2021. Chapter to appear in: *Online and Matching-Based Market Design.* Federico Echenique, Nicole Immorlica and Vijay V. Vazirani, Editors. Cambridge University Press. © 2021. We gratefully acknowledge financial support from the European Research Council under the grant EMBED #757229 (Elliott) and the Cambridge Institute for New Economic Thinking (Talamàs). Felix Mylius and Alastair Langtry provided excellent research assistance.

† Cambridge University.

‡ IESE Business School and CEPR.

Section 7 indicates the core work that we draw on and briefly discusses some closely related questions and areas of active research.

2 Motivating example

Consider a market that matches designers with design studios. Before entering this market, each designer and each design studio can make non-contractible investments—ranging from very specific to very general—that shape the gains from trade in the matching stage.

In the case of designers, general investments might include learning skills that are valuable in any design studio, and specific investments might include learning how to use a particular software that is only used by a particular design studio.

In the case of studios, general investments might include commercial efforts to strengthen connections with potential clients and financiers, and more specific investments might include acquiring hardware, software and office space that facilitate certain kinds of design processes.

We are interested in understanding the extent to which designers and design studios have incentives to invest efficiently. Focusing on transferable utility, this is simply the extent to which the private incentives to make non-contractible investments in such markets are consistent with the maximization of the gains from trade.

The answer to this question hinges on the designers' and studios' expectation of the returns each kind of investment will generate. This depends, in turn, on their expectation of the investments that the others will make as well as how their investment choices will affect who they will match to and on what terms.

This chapter focuses on the following questions: First, how do the bargaining outcomes reached in the matching stage affect the extent to which designers and studios appropriate the marginal returns of their different investments? Second, are there bargaining outcomes or norms that guarantee that all designers and studios have incentives to make efficient investments? Finally, are designers and studios more likely to have good incentives to make general investments than specific investments?

3 Model

The set of agents is $N := I \cup J$, where I is a finite set of workers and J is a finite set of firms. There are two stages. In stage one, agents make non-contractible investments that shape their matching surpluses. In stage two, they bargain over which matches to form and how to share the resulting surplus.

3.1 Stage 1: Investments

Each agent k in N simultaneously chooses an attribute a_k in A_k at cost $c_k(a_k)$. These choices represent non-contractible investments. For instance, in the example above, the possible investments include specific and general investments available to both designers and studios.

The matching surplus generated when worker i matches to firm j depends on their attribute choices, and is given by the function $s : A_i \times A_j \to \mathbb{R}_+$. If an agent does not match with any other agent, we think of her as being matched to herself, and we set $s(a_k, a_k) = 0$ for all a_k in A_k.

For each agent k, there is a null attribute $0 \in A_k$ such that $c_k(0) = 0$, and the surplus that agent k with this null attribute generates with any other agent is 0. Choosing attribute 0 can be interpreted as staying out of the market.

3.2 Stage 2: Pairwise stable outcome

Once the attribute choices have been made, each worker can match with at most one firm, and vice versa. We model this matching process as a cooperative game.

Given the agents' choices of attributes, the matching game in stage 2 is an *assignment game*: There are two populations of agents (workers and firms), with each pair of agents (one from each population) being able to generate weakly positive surplus if they match.

An *outcome* in the assignment game is a matching (each worker matching with no more than one firm and each firm matching to no more than one worker) and a payoff profile. We denote worker i's payoff by $w_i \geq 0$ and firm j's payoff by $v_j \geq 0$.

Definition 1. A *matching* is a function $\mu : N \to N$ such that (i) each worker is mapped either to a firm or to herself, (ii) each firm is mapped either to a worker or to itself, and (iii) worker i is mapped to firm j if and only if firm j is mapped to worker i.

Definition 2. A payoff profile $(w, v) \in \mathbb{R}_+^I \times \mathbb{R}_+^J$ is *feasible for the matching* μ if $w_i = 0$ whenever $\mu(i) = i$, $v_j = 0$ whenever $\mu(j) = j$, and $w_i + v_{\mu(i)} \leq s(a_i, a_{\mu(i)})$ otherwise. A payoff profile is *feasible* if it is feasible for some matching.

Definition 3. A payoff profile (w, v) is *stable* if it is feasible and, for all workers i and all firms j, $w_i + v_j \geq s(a_i, a_j)$. A matching associated with a stable payoff profile is a *stable matching*.

3.3 Stable matchings

Fix an attribute profile $a \in A := \Pi_{k \in N} A_k$. For brevity, for every worker-firm pair (i, j), we denote $s(a_i, a_j)$ by $s(i, j)$. A payoff profile is stable if and only if it is in the core of the assignment game—where the value $V(S)$ of any coalition $S \subseteq N$ is the maximum surplus that can be generated by matching pairs of agents in S. We sometimes write $V(a)$ for $V(N)$.

An immediate and useful implication of stable payoffs being in the core is that in all stable outcomes a surplus maximizing match must be selected (otherwise, the grand coalition would have a profitable deviation). Generically, there is a unique such match, so stability uniquely pins down who matches with whom—but not the resulting payoffs. When there are multiple surplus-maximizing matches, any one of them can support any stable payoff profile. For simplicity, we often assume that there is a unique surplus maximizing match.

A lot is known about the structure of stable payoffs. For example, the set of stable payoff profiles forms a complete lattice both for the partial ordering of workers' payoffs and for the partial ordering of firms' payoffs. In particular, there is a worker-optimal stable payoff profile $(\overline{w}, \underline{v})$ in which every worker simultaneously receives her maximum possible stable payoff and every firm simultaneously receives its minimum possible stable payoff. Similarly, there is a firm-optimal stable payoff profile $(\underline{w}, \overline{v})$ in which every firm simultaneously receives her maximum possible stable payoff and every worker simultaneously receives her minimum possible stable payoff.

3.4 Equilibrium

There are usually many stable outcomes associated with any given attribute profile a. Hence, in order to understand agents' investment incentives in stage 1, we need to specify which stable outcome in stage 2 is associated with each attribute profile chosen in stage 1. We describe this selection via a bargaining function $g : A \to \mathbb{R}_+^N$ that selects a stable payoff profile for each attribute profile a. For any attribute profile a and any agent k, we denote by (a_{-k}, a'_k) the profile a but with a_k swapped for a'_k.

Definition 4. An *equilibrium* is a pair (g, a) such that

(i) $g : A \to \mathbb{R}_+^N$ where, for any attribute profile a, we have that $g(a) = (w(a), v(a))$ is a stable payoff profile in the assignment game with attribute profile a.

(ii) For each worker i and each of her potential attributes a'_i in A_i,

$$w_i(a_{-i}, a_i) - c_i(a_i) \geq w_i(a_{-i}, a'_i) - c_i(a'_i).$$

(iii) For each firm j and each of its potential attributes a'_j in A_j,

$$v_j(a_{-j}, a_j) - c_j(a_j) \geq v_j(a_{-j}, a'_j) - c_j(a'_j).$$

4 Private Investment Incentives

As our focus is on private investment incentives in decentralized markets, we model bargaining over the terms of trade in reduced form through a bargaining function g which is exogenously given (e.g., determined by exogenous institutions and social norms). For example, the sharing of surplus between designers and studios might be determined by historical factors, their different degrees of experience and familiarity with the matching and bargaining process, etc. We will sometimes consider specific bargaining functions.

Example 5. Given an attribute profile a, the *midpoint bargaining function* g is such that, if i and j are matched in the surplus maximizing match, the surplus they generate is split as follows

$$\left(\underline{w}_i(a) + \frac{s(a_i, a_j) - \underline{w}_i(a) - \underline{v}_j(a)}{2},\ \underline{v}_j(a) + \frac{s(a_i, a_j) - \underline{w}_i(a) - \underline{v}_j(a)}{2}\right).$$

where $\underline{w}_i(a)$ and $\underline{v}_j(a)$ are the minimum core payoffs of i and j when the attribute profile is a. In other words, each agent receives her minimum core payoff, and then any remaining surplus in each stable match is split equally.

Since the maximum core payoffs of a worker i and firm j who match under the optimal match are $\overline{w}_i = s(a_i, a_j) - \underline{v}_j$ and $\overline{v}_j = s(a_i, a_j) - \underline{w}_i$, respectively, this payoff profile can be rewritten as

$$\left(\frac{\underline{w}_i + \overline{w}_i}{2}, \frac{\underline{v}_j + \overline{v}_j}{2}\right).$$

In other words, each agent receives the midpoint between her minimum core payoff and her maximum core payoff. As the core is convex, these payoffs are in the core and hence stable.

Given a bargaining function g, we are interested in understanding the extent to which efficient attribute choices can arise in equilibrium regardless of the investment opportunities available on either side.

For concreteness, in this section we focus on the investment incentives of an arbitrary firm j. Given the symmetry between workers and firms, the analysis also holds, *mutatis mutandis*, for an arbitrary worker.

As we are in a transferable utility environment, an attribute profile a is Pareto efficient, henceforth *efficient*, if and only if it maximizes the surplus generated net of investment costs, $V(a) - \sum_{k \in N} c_k(a_k)$. An efficient attribute profile a is also *constrained efficient*, in the sense that each agent k's choice of attribute a_k maximizes $V(a_{-k}, a_k) - c_k(a_k)$ conditional on the others' attributes a_{-k}.

Definition 6. Given an attribute profile a, firm j's investment is *constrained efficient* if, for all a'_j in A_j, $V(a_{-j}, a_j) - c_j(a_j) \geq V(a_{-j}, a'_j) - c_j(a'_j)$.

Efficient attribute profiles are constrained efficient, but constrained efficient attribute profiles need not be efficient. Indeed, even when everyone makes constrained-efficient investments, coordination problems can lead to inefficient investments.

For example, if designers do not expect studios to make substantial investments, they might not have incentives to make substantial investments either, and vice versa, so neither designers nor studios investing much before entering the market might be constrained efficient. We focus mainly on constrained efficiency, abstracting away from coordination problems. (Having said this, we do touch on coordination problems at various junctures.)

Because stable outcomes need not give anyone the full social return of her attribute choices, constrained-efficient attribute choices need not be compatible with private investment incentives. The following result identifies a condition on the bargaining function g that is both necessary and sufficient to guarantee that a given agent makes constrained efficient investments in equilibrium.

Definition 7. Given an attribute profile a, firm j's *ex-post marginal product* is $V(a)-V(a,-j)$, where $V(a,-j)$ denotes the maximum total surplus achievable subject to the constraint that firm j matches to itself.

Proposition 8. *Consider an arbitrary bargaining function g, and an arbitrary firm j with a non-trivial attribute choice (i.e., $|A_j| \geq 2$).*

(i) If, for every attribute profile a, $g_j(a)$ is j's ex-post marginal product, then j's investment is constrained efficient <u>in all equilibria</u>.

(ii) If there exists an attribute profile a such that $g_j(a)$ is not j's ex-post marginal product, then there exist cost functions $\{c_k\}_{k \in N}$ such that j's investment is not constrained efficient <u>in any equilibrium</u>.

Proof of (i). Suppose that the bargaining function g gives firm j its ex-post marginal product; that is,

$$v_j(a') = V(a') - V(a', -j) \quad \text{for every attribute profile } a'. \tag{1}$$

Consider an arbitrary equilibrium attribute profile a. Since j's equilibrium attribute choice is a best response to the others' choices, we have that, for every a'_j in A_j, $v_j(a) - c_j(a_j) \geq v_j(a_{-j}, a'_j) - c_j(a'_j)$ or, using (1),

$$V(a) - V(a, -j) - c_j(a_j) \geq V(a_{-j}, a'_j) - V(a_{-j}, a'_j, -j) - c_j(a'_j).$$

Since $V(a, -j) = V(a_{-j}, a'_j, -j)$, it follows that

$$V(a) - c_j(a_j) \geq V(a_{-j}, a'_j) - c_j(a'_j) \text{ for all } a'_j \text{ in } A_j.$$

That is, j's attribute choice a_j is constrained efficient. □

Proof of (ii). Consider an attribute profile a such that $g_j(a)$ is not j's ex-post marginal product. In particular, suppose that $g_j(a) = V(a) - V(a, -j) - \epsilon$, where $\epsilon \neq 0$. Note that $a_j \neq 0$, since a firm choosing the null attribute 0 necessarily obtains its marginal product 0. Note also that ϵ must be positive: Otherwise, firm j would receive more than its marginal contribution to total surplus, and the coalition of all agents other than j would have a profitable deviation. In other words, such a payoff profile would not be in the core, and hence it would not be stable.

Suppose that, for every attribute $\widehat{a}_j$ in $A_j - \{a_j, 0\}$, $c_j(\widehat{a}_j)$ is sufficiently high so that choosing $\widehat{a}_j$ is a dominated strategy in stage 1. Suppose also that, for every agent $k \neq j$, $c_k(a_k)$ is sufficiently low so that she will choose a_k in equilibrium.[1] Finally, suppose that $c_j(a_j)$ satisfies

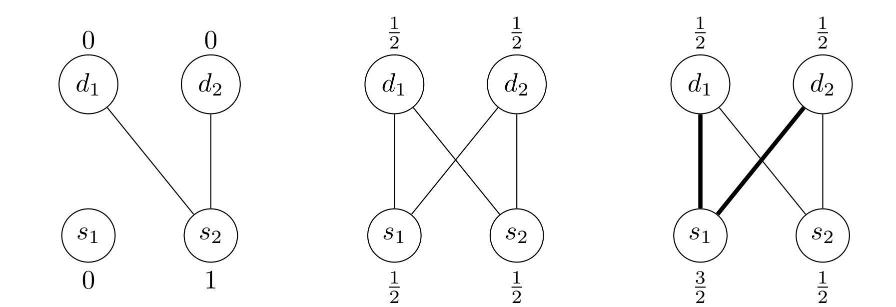
$$0 < V(a) - V(a, -j) - c_j(a_j) < \epsilon.$$

In this case, $(a_{-j}, 0)$ is not constrained efficient because, given the others' attributes a_{-j}, the null attribute 0 generates 0 social surplus, while the attribute a_j generates strictly positive social surplus. But $(a_{-j}, 0)$ is the only attribute profile that can be implemented in equilibrium. Indeed, choosing a_j gives worker j a strictly negative net payoff when the others choose a_{-j}, while her net payoff from choosing 0 is 0. $\square$

The following example illustrates the role that our assumption that firm j has access to the null attribute 0 plays in Part (ii) of Proposition 8: Without this assumption, firm j's equilibrium investment might be constrained-efficient for all cost functions even if she never receives her marginal product.

Example 9. Consider a simple version of the example described in section 2 in which there are only two designers, d_1 and d_2, two design studios, s_1 and s_2, and only general-purpose investments available. Suppose that the bargaining function is the midpoint one described in Example 5. Everyone can choose between attributes 0, 1 and 2, with costs 0, c_1 and c_2, respectively, and the surplus of each match is the product of its members' attributes; that is, $s(a_i, a_j) = a_i a_j$.

Let's fix the attributes of everyone but studio s_1 to be 1, and focus on this studio's private investment incentives. The three possible payoff profiles (corresponding to s_1's three possible attributes) are illustrated in Figure 1.



(a) s_1 picks attribute 0. (b) s_1 picks attribute 1. (c) s_1 picks attribute 2.

Figure 1: Thin and thick links between nodes indicate pairs of agents that generate 1 and 2 units of surplus by matching, respectively. The numbers in each panel represent the stable payoffs chosen by the midpoint bargaining function.

[1] This construction might require setting $c_k(a_k) < 0$, which is a legitimate possibility in our framework. Similar statements can be proved if we were only to allow for non-negative investment costs.

If studio s_1 chooses attribute 0, its ex-post marginal product is 0, and it obtains a gross payoff of 0. If it chooses attribute 1, its ex-post marginal product is 1, and it obtains a gross payoff of 1/2. If it chooses attribute 2, its ex-post marginal product is 2, and it obtains a gross payoff of 3/2.

Note that, even though s_1 only receives its ex-post marginal product when it chooses attribute 0, it receives the full social value of its investment deviation from attribute 1 to attribute 2. As a result, *if it had no access to attribute* 0, it would choose attribute 1 if $c_2 - c_1 > 1$ and attribute 2 if $c_2 - c_1 < 1$. These choices are surplus maximizing. In particular, in this case, its investment would necessarily be constrained efficient regardless of the investment costs c_1 and c_2.

However, since s_1 does have access to the null investment 0, its optimal investment need not be constrained efficient. For example, if $3/2 < c_2 < 2$ and $1/2 < c_1 < 1$, then s_1's privately-optimal investment is the null attribute 0, even though both of the alternative attribute choices 1 and 2 generate strictly more overall surplus net of investment costs.

One interpretation of the availability of the null investment choice is that it introduces an individual rationality (or participation) constraint. Under this interpretation, this example illustrates how satisfying the incentive-compatibility constraints between the different investment alternatives available to those participating in the market does not guarantee constrained-efficient investments once the participation constraints are taken into account.

Another potential source of inefficiencies is miscoordination. For instance, in this example, inefficient equilibria in which everyone chooses attribute 0 can coexist with equilibria in which everyone obtains strictly positive net payoffs.

5 Efficient investments

Proposition 8 above highlights that giving an agent her ex-post marginal product is crucial to guarantee that she makes constrained-efficient investments in equilibrium. Before discussing the extent to which this property can simultaneously hold for all agents, we provide a useful interpretation of an agent's ex-post marginal product in terms of the value of rematching when she leaves the market.

5.1 Marginal product and the value of rematching

For each subset $S \subseteq N$ of agents, consider an efficient matching $\eta^S : S \to S$. Such a matching η^S must maximize the surplus $V(S)$ obtainable by the coalition S. For brevity, for every agent k in S, we denote $\eta^S(k)$ by η^S_k. Note that, for any firm j,

$$\begin{aligned} V(N) - V(N-j) &= \sum_{j' \in J} s(\eta^N_{j'}, j') - \sum_{j' \in J-j} s(\eta^{N-j}_{j'}, j') \\ &= s(\eta^N_j, j) - \Phi(j) \end{aligned} \tag{2}$$

where $\Phi(j)$ denotes *the value of rematching* (i.e., implementing the new optimal matching η^{N-j}) *after j has been removed from the market.*[2] That is,

$$\Phi(j) := \sum_{j' \in J-j} \left[s(\eta_{j'}^{N-j}, j') - s(\eta_{j'}^{N}, j') \right].$$

The analogous expression holds for any worker i.

5.2 Main result

The following result highlights that there is always a stable payoff profile that gives all the agents on one side of the market their ex-post marginal products.

Theorem 10. *The ex-post marginal product $V(N) - V(N-j)$ of each firm j is its payoff $\overline{v}_j$ in the firm-optimal stable profile.*

The combination of Proposition 8 and Theorem 10 implies that there exists a bargaining function g that ensures that all firms simultaneously have appropriate investment incentives (i.e., the bargaining function g that, for any given attribute profile a, selects the firm-optimal stable payoff profile). An analogous result holds for the worker-optimal stable payoff profile, so there also exists a bargaining function that ensures that all workers have appropriate investment incentives.

Unfortunately, however, the bargaining function that ensures that all firms simultaneously have appropriate investment incentives is different from the bargaining function that ensures that all workers simultaneously have appropriate investment incentives—except in the special case in which the core is a singleton for every possible attribute profile a. Hence, except in this special case, the bargaining function g cannot simultaneously give everyone her marginal product.

Corollary 11. *Unless the core uniquely pins down payoffs for every attribute profile, no bargaining function g can guarantee that private incentives to invest are aligned with the maximization of the gains from trade.*

Given a constrained-efficient attribute profile, the requirement to guarantee that it can be implemented in equilibrium is that its associated assignment game has a singleton core. Indeed, in this case, this attribute profile can be implemented in equilibrium regardless of the investment technology. Unfortunately, however, there is no reason to expect this to necessarily be the case in applications—even as markets become thick.

In the case of exogenous attributes, an active area of study investigates the conditions on the matching surpluses under which the core shrinks to a point when the market becomes large. In the presence of investment decisions, however, the matching surpluses are endogenous, and there's no reason to expect the endogenous attribute choices to lead to a small core—even as the market gets large.

[2]For any set S, we often write $S - j$ (rather than $S \setminus \{j\}$) to mean S excluding j.

For any given attribute profile, Proposition 12 below provides three ways to express workers' minimum stable payoffs in terms of the surpluses that are available. These expressions are helpful for understanding investment incentives in the typical scenario in which the core does not uniquely pin down outcomes uniquely.

Proposition 12. *In the firm-optimal stable payoff profile, worker i's payoff is*

(i) the value $\Phi(\eta_i^N)$ of rematching after her partner η_i^N leaves the market,

(ii) her ex-post marginal product $V(N - \eta_i^N) - V(N - \eta_i^N - i)$ when the set of agents is everyone but her partner η_i^N, and

(iii) her ex-post marginal product when an identical agent enters the market.

An immediate implication of Proposition 12 is that a worker i receives her ex-post marginal product at the firm-optimal stable payoff profile if and only if her ex-post marginal product is the same before and after her partner η_i^N in the optimal matching leaves the market.

We defer the proofs of Theorem 10 and Proposition 12 to Section 6. We now discuss some of their main implications for investment incentives.

5.3 Generality of the investment technology

It is worth noting that we have placed very little structure on the investment technology. This makes the results widely applicable. For instance, as in Example 9, the framework can capture investments on the extensive margin (e.g., costly participation decisions by designers to be active in the labor market and by studios to create vacancies) as well as on the intensive margin (e.g., how many years of education a designer pursues and how much hardware and software a studio acquires). In addition, there can be investments that are completely specific and only increase the surplus of a particular match, other investments that are general and increase the surplus of every match, and the whole spectrum of possibilities in between.

5.4 One-sided investments

A special case of interest is the one in which only one side of the market has non-contractible investment choices before entering the market. This would be the case if, for example, studios had opportunities to invest prior to matching but designers did not. In this case, an immediate implication of Proposition 8 and Theorem 10 is that, if g selects the studio-optimal core payoff profile then the investments in stage 1 will necessarily be constrained efficient in equilibrium.

On the one hand, this is a beautifully sharp result that holds for any investment technology that the agents on the investing side of the market might have access to. On the other hand, the practical relevance of this result depends on the extent to which we expect the process of bargaining and matching to result in this particular extreme point of the core.

We might hope that social norms and institutions have a tendency to evolve to enhance efficiency. This explanation gains credence once we note that there are simple natural mechanisms that lead to the required bargaining function. Indeed, if a social norm implies that workers can only increase their wages using outside offers, then a corresponding dynamic adjustment process that starts with very low wages will bid up wages only until the firm-optimal stable payoffs are obtained. The key behind these wage adjustment mechanisms is that there is an asymmetry in terms of which side of the market makes the offers.

5.5 Two-sided under and over investment

When both sides of the market have access to investments, it is typically impossible to simultaneously provide appropriate investment incentives to everyone at the same time. Proposition 12 can help us understand the inefficiencies that are likely to be present in this case.

For the sake of illustration, suppose that the bargaining function implements the firm-optimal payoffs, but that both workers and firms have investments to make before entering the market. We have already seen that workers are then likely to have poor incentives to make investments that increase the value of their match. Indeed, worker i's payoff is equal to the value of rematching when i's partner η_i^N leaves the market, which need not be related to $s(i, \eta_i^N)$. Hence worker i will, at the margin, not have any incentive to make an investment that only increases the surplus $s(i, \eta_i^N)$ that she generates with the person with whom she matches in equilibrium.

This case is a particularly acute example of the hold-up problem in matching markets. But the troubles do not only come from potential underinvestment: The fact that, in the firm's optimal stable payoff profile, worker i's payoff is equal to the value $\Phi(\eta_i^N)$ of rematching implies that worker i only has incentives to pursue investments that increase the surplus of matches that do not occur in equilibrium.

In particular, worker i might not only have incentives to underinvest in the match that she will be part of in equilibrium, but also *to overinvest in alternative matches that she will not be part of in equilibrium.*

Investments that only increase the surplus of unrealized matches are pure overinvestments—in the sense that they generate no new surplus and serve only to redistribute existing surplus. Moreover, while these problems can be ameliorated by moving away from the extreme point of the core, unfortunately this only shifts these problems to the other side of the market.

5.6 General-purpose investments

Investment inefficiencies are particularly acute when investments are purely relationship specific. When worker i's investments are instead of a more general-purpose nature, they might increase the value $\Phi(\eta_i^N)$ of rematching at a similar rate as the surplus $s(i, \eta_i^N)$ of her equilibrium match. If worker i's available investments change the value $\Phi(\eta_i^N)$ by the same amount as the surplus $s(i, \eta_i^N)$

of her equilibrium match, then Proposition 12 implies that worker i fully appropriates the social rewards of her marginal investment deviations. Hence, even if the firm-optimal payoffs are selected, when the relevant investments are general purpose, workers can have appropriate investment incentives as well.

6 Proofs of the main results

To prove Theorem 10 and Proposition 12 above, it is helpful to first put some structure on how an agent's marginal product depends on the available matching opportunities. Let I_j and J_j be the set of workers and firms that rematch when an arbitrary firm j is removed from the market; that is,

$$I_j := \{i \in I : \eta_i^N \neq \eta_i^{N-j}\} \text{ and } J_j := \{j' \in J : \eta_{j'}^N \neq \eta_{j'}^{N-j}\}.$$

Also let $N_j := I_j \cup J_j$, and denote the cardinality of this set by n_j. The following Lemma describes useful properties of the optimal rematchings that occur when an arbitrary firm j is removed from the market.

Lemma 13. *For every firm j, consider the sequence $\mathcal{S}_j = \mathcal{S}_j(1), \ldots, \mathcal{S}_j(n_j)$ with $\mathcal{S}_j(1) = j$ and, for every $1 \leq q < n_j$,*

(i) if $\mathcal{S}_j(q)$ is in J, then $\mathcal{S}_j(q+1) = \eta_{\mathcal{S}_j(q)}^N$;

(ii) if $\mathcal{S}_j(q)$ is in I, then $\mathcal{S}_j(q+1) = \eta_{\mathcal{S}_j(q)}^{N-j}$.

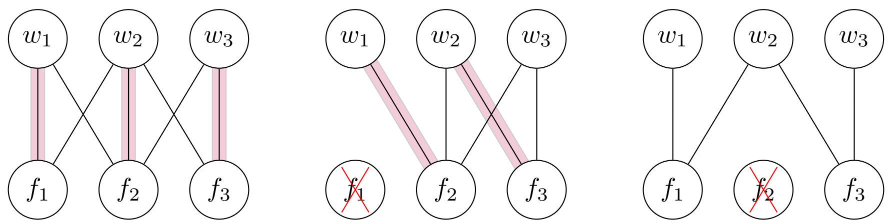
The sequence $\mathcal{S}_j$ contains every agent in N_j. Moreover, for any $1 \leq q \leq n_j$, if $\mathcal{S}_j(q)$ is in J, then $\mathcal{S}_{\mathcal{S}_j(q)} = (\mathcal{S}_j(q'))_{q'=q}^{n_j}$.

In other words, the optimal sequence of re-matchings after the removal of a firm j has a chain-like structure that is such that, if we cut it at any firm $\mathcal{S}(q)$, the remaining sequence $\mathcal{S}(q), \mathcal{S}(q+1), \ldots, \mathcal{S}(n_i)$ determines the optimal rematching when this firm $\mathcal{S}(q)$ is removed from the market instead of j. We leave the proof of Lemma 13 as an exercise, and we illustrate it here with a simple example.

Example 14. Suppose that the matchings illustrated in Figures 2a and 2b are the only optimal ones given that the set of agents is $\{w_1, w_2, w_3, f_1, f_2, f_3\}$ and $\{w_1, w_2, w_3, f_2, f_3\}$, respectively.

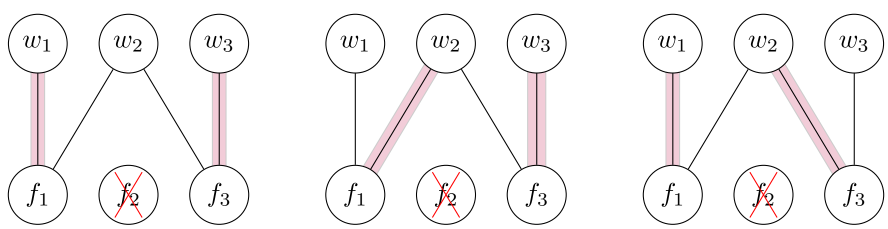
Note that all the remaining workers and firms are involved in the optimal rematching after firm f_1 is removed from the market, and the sequence $\mathcal{S}_{f_1}$ defined by Lemma 13 is $(f_1, w_1, f_2, w_2, f_3, w_3)$. Consider the case in which firm f_2 is removed instead of firm f_1. Lemma 13 implies that the optimal rematching sequence $\mathcal{S}_{f_2}$ in this case (f_2, w_2, f_3, w_3).

Consider an optimal matching when the agent set is $\{w_1, w_2, w_3, f_1, f_3\}$ (see Figure 2c). Clearly, this must be one of the three rematchings illustrated in Figure 3. Lemma 13 implies that the surplus-maximizing matching is the one



(a) Full market (b) Removing f_1 (c) Removing f_2

Figure 2: Thin links between nodes indicate pairs of agents that generate strictly positive surplus by matching. In panels (a) and (b), thick purple shaded links between nodes identify the surplus maximizing matches. In panel (c), the surplus maximizing matches are not shown.



(a) Rematch A. (b) Rematch B. (c) Rematch C.

Figure 3: Thin links between nodes indicate pairs of agents that generate strictly positive surplus by matching. Thick purple shaded links between nodes identify candidate surplus-maximizing matchings.

shown in Figure 3c. To see that this is indeed the surplus-maximizing matching, note that

$$s(w_1, f_1) + s(w_2, f_3) > \max\left[s(w_1, f_1) + s(w_3, f_3), s(w_2, f_1) + s(w_3, f_3)\right].$$

The fact that $s(w_1, f_1)+s(w_2, f_3) > s(w_1, f_1)+s(w_3, f_3)$ follows directly from the fact that the surplus-maximizing matching depicted in Figure 2(b) is optimal. The fact that $s(w_1, f_1) + s(w_2, f_3) > s(w_2, f_1) + s(w_3, f_3)$ follows from the fact that, since the matching depicted in Figure 2a is optimal, $s(w_1, f_1)+s(w_2, f_2) > s(w_1, f_2) + s(w_2, f_1)$ and, since the matching depicted in Figure 2b is optimal, $s(w_1, f_2) + s(w_2, f_3) > s(w_2, f_2) + s(w_3, f_3)$.

Proof of Theorem 10. We show by construction that there are stable payoffs in which $v_j = V(N) - V(N - j)$ for all firms j. These are the firms' optimal stable payoffs because, if $v_j > V(N) - V(N - j)$, then the coalition $N - j$ would have a profitable deviation.

Let $v_j = V(N) - V(N - j)$ *for all firms* j. If worker i matches with a firm j in the efficient match η^N, set $w_i = s(i, j) - v_j$; otherwise set $w_i = 0$. It is enough to show that these payoffs (w, v) are stable.

For every worker i, $w_i = s(i, \eta_i^N) - v_{\eta_i^N}$ (see Theorem [[7, part 1]] in Chapter [[**??**]]).[3] Moereover, our choice of stable payoff profile implies that

$$w_i = s(i, \eta_i^N) - v_{\eta_i^N} = s(i, \eta_i^N) - (V(N) - V(N - \eta_i^N)) = \Phi(\eta_i^N).$$

Hence, the most worker i can receive by deviating to match with firm $j \neq \eta_i^N$ is

$$s(i, j) - v_j = s(i, j) - \left[s(\eta_j^N, j) - \Phi(j)\right].$$

It is enough to show that

$$\Phi(\eta_i^N) \geq s(i, j) - s(\eta_j^N, j) + \Phi(j) \text{ for all firms } j. \tag{3}$$

That is, the value of rematching when i's partner η_i^N leaves the market must be at least as large as the net surplus $s(i, j) - s(\eta_j^N, j)$ generated by matching i with any firm j plus the value $\Phi(j)$ of rematching when j leaves the market.

This is easiest to see in the case in which worker i is not in the set of workers that must rematch for the gains from trade to be maximized when j is removed (i.e., $i \notin I_j$). In this case, the rematching among the agents in N_j that generates rematching value $\Phi(j)$ when firm j is removed is feasible after worker i deviates to match to j. Hence, when η_i^N is removed from the market, this rematching along with the matching of i to j is feasible. As $\Phi(\eta_i^N)$ is the value of rematching when η_i^N is removed from the market, inequality (3) must hold in this case.

To derive inequality (3) in the alternative case in which worker i is in I_j, we rely on Lemma 13: The agents in N_j can be arranged into a sequence $\mathcal{S}_j$

[3] To prevent i and η_i^N having a jointly profitable deviation, we need $w_i + v_{\eta_i^N} \geq s(i, \eta_i^N)$; and to prevent the other agents having a profitable coalitional deviation, we need $w_i + v_{\eta_i^N} \leq s(i, \eta_i^N)$.

with $\mathcal{S}_j(q) = \eta_i^N$ for some $1 \le q \le n_j$, and $(\mathcal{S}_j(q'))_{q'=q}^n = \mathcal{S}_{\eta_i^N}$. Letting $\widehat{J}_j := J_j \cap \{\mathcal{S}_j(1), \ldots, \mathcal{S}_j(q)\}$, we have that

$$\begin{aligned}\Phi(j) &= \sum_{j' \in J_j - j} \left[s(\eta_{j'}^{N-j}, j') - s(\eta_{j'}^N, j') \right] \\ &= \underbrace{\sum_{j' \in \widehat{J}_j - j} \left[s(\eta_{j'}^{N-j}, j') - s(\eta_{j'}^N, j') \right]}_{\Psi(\eta_i^N, j)} + \Phi(\eta_i^N)\end{aligned}$$

Hence, (3) follows from the fact that $\Psi(\eta_i^N, j)$ is bounded above by $s(\eta_j^N, j) - s(i, j)$. Indeed, otherwise it would be possible to match the firms in $\widehat{J}_j - j$ according to η^{N-j}, and worker i to firm j, while matching all other workers according to η^N, and this would generate more surplus than matching all workers according to η^N, contradicting the optimality of the match η^N. □

Proof of Proposition 12. For worker i, as argued in the proof of Theorem 10, we must have $w_i = s(i, \eta_i^N) - v_j$ in any stable outcome with all agents present. Thus, we must have that

$$\underline{w}_i = s(i, \eta_i^N) - \overline{v}_{\eta_i^N} = s(i, \eta_i^N) - \left(V(N) - V(N - \eta_i^N) \right) = \Phi(\eta_i^N),$$

where the second equality follows from Theorem 10 and the final equality follows from equation (2).

For part (ii), we need to show that

$$\underline{w}_i = V(N - \eta_i^N) - V(N - \eta_i^N - i). \tag{4}$$

First note that

$$\underline{w}_i = s(i, \eta_i^N) - \overline{v}_{\eta_i^N} = s(i, \eta_i^N) - V(N) + V(N - \eta_i^N)$$

for all workers i. Equation (4) then follows from the fact that, because when an efficiently matched pair of agents (in this case i and η_i^N) are both removed, the surplus maximizing match leaves all other agents matched as they were before, so $V(N) = s(i, \eta_i^N) + V(N - i - \eta_i^N)$.

We leave part (iii) as an exercise.

□

7 Discussion and bibliographical notes

7.1 Investment efficiency and strategy proofness

The main result of this chapter, Theorem 10, is due to Leonard (1983) (see also Demange (1982)). This seminal paper considers a mechanism designer who wants to achieve efficiency while eliciting workers' preferences. He shows that the worker-optimal core allocation gives every worker her marginal product, so the mechanism designer can elicit workers' preferences and obtain efficiency by selecting this point in the core.

For a very nice and well-known application of the key ideas in this chapter in the context of one-sided investments (in this case, to form relationships that facilitate trade), see Kranton and Minehart (2001). The differences between agents' incentives to make general vs. specific investments were pointed out by Becker (1962) and, remarkably, the main insights that we discuss regarding these differences were already anticipated there.

Our claim that the firm-optimal stable outcome arises when workers rely on outside offers to bid up their wages is broadly based on the dynamic adjustment process described in Crawford and Knoer (1981). This process is generalized in Kelso and Crawford (1982) and Demange et al. (1986).[4]

The connection between strategy-proofness and investment efficiency goes back at least to Rogerson (1992), who shows that the Vickrey-Clarke-Groves mechanisms induce efficient investments because they force each agent to internalize the social gains or losses from changes in his or her valuation over outcomes.

The connection between giving workers good incentives to truthfully report their type and implementing the worker-optimal core outcome also holds in non-transferable utility (NTU) environments. In such environments, when workers submit preference lists and the deferred acceptance algorithm allocates them to firms, the deferred acceptance algorithm is dominant strategy incentive compatible and implements the worker-optimal match. However, in NTU environments there is not an equivalence between creating good incentives for truthfully revealing types and creating efficient investment incentives.

To see this, consider a version of our environment in which the surplus of each match is always split equally between the corresponding worker and firm. This turns our model into an NTU environment. Now suppose that, in the worker-optimal match, worker 1 is matched to firm 1, worker 1 receives wage x and hence firm 1 receives a payoff of x as well. Suppose also that worker 1 has access to an investment that would increase these payoffs by y (to $x + y$), while leaving the payoffs obtained in other matches unaffected. After making such an investment, the worker-optimal match must still match worker 1 to firm 1. Thus, worker 1's payoff will increase by $y/2$. However, if the cost of the investment is $c \in (y/2, y)$ then worker 1 will not make the investment despite it increasing the joint surplus of firm 1 and worker 1.

The key difference between the non-transferable and transferable utility environments is of course the margin on which outcomes can adjust. Whereas in a transferable utility world wages are free to adjust, in an NTU world adjustments can only occur on the extensive margin of who matches to whom. In general then, it is unsurprising that the coarse instrument of adjusting the match is insufficient to provide good investment incentives for investments that operate

[4]The famous Hungarian method developed by Kuhn (1955) also yields this outcome.

on the intensive margin.[5]

7.2 Perfect competition

Gretsky et al. (1999) provide a rigorous formulation of perfect competition in the assignment model as the ability of agents to extract their marginal products or, equivalently, the inability of agents to favourably manipulate prices. They show that imperfect competition is generic in finite assignment economies and perfect competition is generic in a natural family of continuum assignment economies. They provide an elegant but less constructive proof of Theorem 10 and Proposition 12 using the subdifferential of the total surplus function (see their Proposition 1). Our analysis of two-sided over and under investment in finite assignment games is based on Elliott (2015).

7.3 Alternative approach: Bargaining function can depend on investment costs

Taking the bargaining function g as an exogenous object determined by social norms and institutions, we have investigated the extent to which this function gives agents incentives to invest efficiently. In particular, we have focused on the question of when efficiency can be achieved for a given bargaining function g independently of the investment costs. Assuming that the investment choice of agent i includes the option of not entering the market (which imposes a natural individual rationality constraint), we have argued that this requires agent i to obtain her ex-post marginal product in the matching stage.

As we have illustrated in Example 9, when no individual rationality constraints are present, guaranteeing that agent i's equilibrium investments are constrained efficient requires the weaker condition that she obtains the full change in social surplus created by any of her possible investment deviations. Cole et al. (2001) discuss richer examples under which this is the case, and describe conditions that ensure that this weaker property is satisfied in finite markets.

An alternative approach is to ask a related but substantially different question: Given the investment costs, is there a bargaining outcome that ensures that efficient investments can be implemented in equilibrium? This approach does not take the bargaining function as given, but instead asks whether efficient investments are consistent with equilibrium. Cole et al. (2001) show that the answer is yes.

For some intuition, consider the case in which agents can contract on their investments. The resulting investment and bargaining game can be seen as an assignment game, so efficiency (including in this case investment efficiency) can be achieved without any transfers at the investment stage. In particular, if a matching is efficient it is stable in this assignment game with complete contracts. Given that the ability to write contracts only expands the set of

[5]Interestingly, in large markets, adjustments on the extensive margin become richer and it is sometimes possible to provide good investment incentives in NTU environments (see, for example, Peters and Siow (2002)).

potential deviations that agents can entertain, this efficient outcome is also an equilibrium when agents cannot write such contracts.[6]

8 Conclusion

This chapter has emphasized that social norms and institutions that shape how surplus is shared in finite matching markets are unlikely to prevent hold-up problems. We conclude by briefly discussing the sensitivity of this message to the following two important assumptions that we have made throughout: (i) the bargaining function selects a stable payoff profile, and (ii) the market is static—in the sense that there is no entry of new agents over time.

Regarding stability, it seems intuitive that two agents with a profitable pairwise deviation should be able to realize these gains. However, as emphasized by Elliott and Nava (2019), extending standard non-cooperative bargaining models to a market setting yields outcomes that need not be stable even as the non-cooperative bargaining frictions vanish, and this has the potential to further distort agents' investment incentives.

Regarding entry, in the context of a general non-cooperative bargaining game with dynamic entry, Elliott and Talamàs (2021) show that the hold-up problem necessarily disappears as bargaining frictions vanish if and only if there is a minimal amount of competition always present in the market. This suggests that taking into account the dynamic nature of many matching markets of interest can overturn some of the negative implications of the results highlighted in this chapter.

9 Exercises

1. Suppose that there are two workers and two firms. Each can choose either to *invest*, or to *not invest*. The matching surplus of each worker-firm pair is

2	units of surplus	if *both* have invested,
1	unit of surplus	if *only one of them* has invested, and
0	units of surplus	if *none of them* has invested.

Not investing costs zero, and investing costs c, with $1/2 < c < 1$. Hence, efficiency requires that everyone invests.

(A) Show that a matching is efficient unless it matches a worker to herself.

(B) Let $x_i \in \{0, 1\}$ represent agent i's investment choice. Find a function $g : \{0, 1\} \to \mathbb{R}_+$ such that, if each agent i that has invested x_i receives payoff $g(x_i)$, then these payoffs are feasible (sum to weakly less than the overall surplus obtained) and agents' private investment incentives are aligned with social investment incentives—i.e., for every agent i and any

[6]See Nöldeke and Samuelson (2015) for a detailed exposition of this argument.

investments of the other agents, agent i finds it optimal to invest if and only if this increases the total net surplus generated by an efficient match.

(C) Show that there is a stable outcome where each agent i's payoff is $g(x_i)$.

(D) Let $S \in \{0,1,2\}^4$ be a tuple describing the surplus that can be obtained by any worker-firm pair. Show that there *does not* exist any function $f : \{0,1,2\}^4 \to \mathbb{R}^4_+$ such that the payoffs $f(S)$ are feasible and private investment incentives are always aligned with social investment incentives.

2. Consider n workers and $m < n$ firms. Each worker i has a different type $\alpha_i \in [0,1]$, and each firm j also has a different type $\beta_j \in [0,1]$. We label workers 1 to n and firms 1 to m so that $\alpha_1 > \alpha_2 > \cdots > \alpha_n$ and $\beta_1 > \beta_2 > \cdots > \beta_m$. The matching surplus of worker i and firm j is $\alpha_i\beta_j$. The bargaining function selects the firms' optimal stable payoffs. We say that a matching is *assortative* if, for $k = 1, \ldots, m$, firm k matches to worker k.

(A) Show that a stable matching is necessarily assortative.

(B) Derive the firms' optimal stable payoffs.

(C) Suppose that, prior to matching, an arbitrary firm k can make an investment that increases its matching surplus with worker i to $\alpha_i\beta_k + r$ where $r > 0$, while leaving the other surpluses unaffected. The cost of such an investment is $c > 0$. Should this firm invest?

(D) Now suppose instead that, prior to matching, every worker i has access to an investment $x_i \in [0,1]$ that improves her type. Specifically, each worker i can increases her type to $\theta_i + r(x_i)$, where r is a strictly increasing, twice differentiable and concave function with $r(0) = 0$. The cost of worker i making an investment x_i is $c(x_i)$, where c is an increasing, twice differentiable and strictly convex function with $c(0) = 0$, $c'(0) = 0$ and $\lim_{x\to 1} = \infty$.

   (i) Assuming assortative matching, find the payoff maximizing investment of worker m.

   (ii) Assuming assortative matching, find the optimal investment of each worker.

   (iii) Show that there is an equilibrium with assortative matching.

   (iv) Show that the investments in this equilibrium are inefficient.

   (v) Let S and $S^\star$ be the total net surplus generated by efficient investments and equilibrium investments, respectively. Letting $\Delta = \beta_1 - \beta_m$, show that $\lim_{\Delta\to 0}(S - S^\star) > 0$.

   (vi) Suppose instead that $n < m$, and consider an equilibrium in which each worker k makes an investment $r'(x_k)\beta_{k+1} = c'(x_k)$. Show that $\lim_{\Delta\to 0}(S - S^\star) = 0$.

3 (Challenging)**.** In this exercise you are asked to extend the logic of Example 14 to prove Lemma 13.

(A) Show that Lemma 13 holds trivially if $\eta_i^N = i$.

(B) Show that a weak subset of the agents in N_i can be arranged into a sequence starting with agent η_i^N, satisfying

  (a) if $\mathcal{S}_i(q)$ is in I, then $\mathcal{S}_i(q+1) = \eta^N_{\mathcal{S}_i(q)}$

  (b) if $\mathcal{S}_i(q)$ is in J, then $\mathcal{S}_i(q+1) = \eta^{N-i}_{\mathcal{S}_i(q)}$,

that either (i) ends once it reaches either a firm $j \in J_i$ such that $\eta_k^{N-i} = j$; (ii) ends once it reaches a worker $k \in I_i$ such that $\eta_k^N = k$; or (iii) does not end.

(C) Show that the above sequence must end.

(D) Show that all agents in N_i can be arranged into a sequence like that described in part (B).

(E) Show that, if $\mathcal{S}_i(q') = k \in I$, then $N_k = \{\mathcal{S}_i(q) : q \geq q'\}$.

4 (Challenging)**.** This exercise guides you through the proof of part (iii) of Proposition 12.

(A) Let i' be an identical copy of i. Show that if $\eta_i^{N+i'} = \eta_i^N$ or $\eta_{i'}^{N+i'} = \eta_i^N$ then, in the firm-optimal stable payoff profile (when the set of agents is N), worker i's payoff is its ex-post marginal product after i' has been added to the market.

(B) Show that either $\eta_i^{N+i'} = \eta_i^N$ or $\eta_{i'}^{N+i'} = \eta_i^N$.

References

Becker, Gary S. 1962. Investment in human capital: A theoretical analysis. *Journal of Political Economy*, **70**(5, Part 2), 9–49.

Cole, Harold, Mailath, George, and Postlewaite, Andrew. 2001. Efficient non-contractible investments in finite economies. *The BE Journal of Theoretical Economics*, **1**(1), 1–34.

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

Demange, Gabrielle. 1982. Strategyproofness in the Assignment Market Game. *Labratorie d' Econométrie de V Ecole Polytec.*

Demange, Gabrielle, Gale, David, and Sotomayor, Marilda. 1986. Multi-item auctions. *Journal of Political Economy*, **94**(4), 863–872.

Elliott, Matt, and Nava, Francesco. 2019. Decentralized bargaining in matching markets: Efficient stationary equilibria and the core. *Theoretical Economics*, **14**, 211–251.

Elliott, Matthew. 2015. Inefficiencies in networked markets. *American Economic Journal: Microeconomics*, **7**(4), 43–82.

Elliott, Matthew, and Talamàs, Eduard. 2021. Bargaining foundations for price taking in matching markets. *Mimeo*.

Gretsky, Neil E, Ostroy, Joseph M, and Zame, William R. 1999. Perfect competition in the continuous assignment model. *Journal of Economic Theory*, **88**(1), 60–118.

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

Kranton, Rachel E, and Minehart, Deborah F. 2001. A theory of buyer-seller networks. *The American Economic Review*, **91**(3), 485–508.

Kuhn, Harold W. 1955. The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, **2**(1-2), 83–97.

Leonard, Herman B. 1983. Elicitation of honest preferences for the assignment of individuals to positions. *Journal of Political Economy*, **91**(3), 461–479.

Nöldeke, Georg, and Samuelson, Larry. 2015. Investment and competitive matching. *Econometrica*, **83**(3), 835–896.

Peters, Michael, and Siow, Aloysius. 2002. Competing premarital investments. *Journal of Political Economy*, **110**(3), 592–608.

Rogerson, William P. 1992. Contractual solutions to the hold-up problem. *The Review of Economic Studies*, **59**(4), 777–793.