January 25, 2016

Abstract

This paper formulates and analyzes a dynamic assignment model with one-sided adverse selection (unobserved worker characteristics) and moral hazard (unobserved worker effort). It defines a notion of stationary equilibrium in which workers are matched to tasks endogenously on the basis of observable output. For each given payment schedule, such an equilibrium exists and is unique. At equilibrium, adverse selection is eliminated and moral hazard is mitigated. Firm profit in equilibrium is compared against natural benchmarks.

JEL classification TO BE ADDED

Key Words: TO BE ADDED

Endogenous Matching in a Dynamic Assignment Model with Adverse Selection and Moral Hazard*

Mihaela van der Schaar[†] Yuanzhang Xiao[‡]

William R. Zame[§]

January 25, 2016

*We are grateful for comments from Marek Pycia and form audiences at EIEF (Rome), IHS (Vienna), UCLA, UCSD and the Network Economics and Game Theory Conference 2014. We acknowledge financial support from the National Science Foundation under grants XXXX. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of any funding agency.

[†]Department of Electrical Engineering, UCLA, Los Angeles, CA 90095

[‡]Department of Electrical Engineering, Northwestern Univeristy, Evanston, IL 60208

[§]Department of Economics, UCLA, Los Angeles, CA 90095; zame econ.ucla.edu (corresponding author)

1 Introduction

The seminal work of Shapley and Shubik (1971) on the assignment model has given rise to a vast literature on models of matching with transfers.[1] These models have found an enormous range of applications: to marriage markets (Becker, 1974), to labor markets (Shimer and Smith, 2000), to international trade (Grossman, Helpman and Kircher, 2014), to perfect competition (Gretsky, Ostroy and Zame, 1999), to hold-ups (Cole, Mailath and Postlewaite, 2001; Makowski, 2004). Most of that literature treats a static, rather than dynamic, environment and focuses on (the implications of) the "proper" matching of the two sides of the market – buyers to sellers, workers to firms, men to women, or, as is in our setting, workers to tasks – and on the division of output/value. This literature largely ignores the possibility that characteristics might not be observable – so that there is a problem of adverse selection. Perhaps more remarkably, it entirely ignores the fact that output does not simply fall from the sky, but must be produced, and that producing output requires costly effort. Thus, it is not enough to match the correct worker to the correct task; it is also necessary to provide incentives for the worker to exert effort – so there is also a problem of moral hazard.

This paper takes some steps toward integrating adverse selection and moral hazard into a dynamic version of the assignment model. We show that in an environment in which interactions are ongoing (and in the presence of natural assumptions), one sided adverse selection can be eliminated and moral hazard can be mitigated in an equilibrium in which the match-

[1] Of course, the seminal work of Gale and Shapley (1962) has given rise to a different but equally vast literature on models of matching *without* transfers.

ing between workers and tasks is determined *endogenously* on the basis of observable output.

Specifically, we consider an infinite horizon discrete time environment with a single long-lived firm and a continuum of long-lived workers W. In each period, a continuum of tasks T arrives to the firm, which matches tasks to workers. If worker w is matched to task t and exerts effort e it produces output $y = Y(e, w, t)$, receives a payment $P(y)$ (according to a fixed and pre-determined payment schedule) and incurs a cost $C(e, w)$, and so receives the net utility $U(e, w, t) = P(y) - C(e, w)$; the firm earns the net profit $y - P(y)$.[2] The ranking of tasks is commonly observed but characteristics of workers (production and cost functions), and of course effort, are not observed. Thus there is both adverse selection and moral hazard (on one side). As is common in the literature, we assume that worker and task types are one-dimensional. We treat the simplest endogenous matching rule: in each period, workers are assigned to tasks according to the ranking of their production in the previous period. For simplicity and tractability we assume that output and cost are multiplicatively separable in effort, worker type and task type and restrict to linear payment schedules. (The most obvious and important implication of separability is that output is supermodular in each pair of effort, worker type and task type. Linear payment schedules ease the comparison of firm profit in our equilibrium and in various benchmarks.

[2]Note that output depends on effort, worker type and task type, and cost depends on effort and worker type, but payment depends only on output; this is in keeping with our intent that the firm observes output but does not observe effort or the production function of workers.

As we discuss below, our assumptions could be weakened – but we prefer to make stronger assumptions in the interests of focusing more on the economic forces in operation and less on technical details.) In this setting, we show that there is a (unique) stationary equilibrium in which matching is perfectly assortative (better workers are matched to better tasks) and (in the presence of an additional assumption on cost) that the firm's profit in the stationary assortative equilibrium is strictly greater than in either of two benchmarks: random matching of workers to tasks, with workers choosing effort optimally, or full information assortative matching of workers to tasks, with workers choosing effort optimally. Thus, the endogenous matching rule – that workers who produce more output are matched with better tasks – leads both to "proper" matching of workers to tasks and to stronger incentives for workers to exert effort. The firm benefits because it is able to provide incentives to workers *both* by paying for output *and* by conditioning future assignments of tasks on current output, so that workers who produce more in the current period are matched to better tasks in the next period.

Our story might be interpreted as a two-stage game: in the first stage the firm chooses and commits to a payment schedule (the firm always matches workers to tasks assortatively by output); in the second stage the workers play a stationary equilibrium given the payment schedule to which the firm has committed. Our focus here is on the second stage; i.e. solving for behavior in the second stage – stationary assortative equilibrium – *given* the matching rule and the payment schedule.

Because it seems an important feature, we emphasize that the informational requirements for stationary assortative equilibrium are remarkably weak. Workers – who might be viewed as the active participants – must observe the ranking of tasks, the output distribution and their own characteristics but not the characteristics of other workers; the firm must observe the ranking of tasks and the output of workers, but not the characteristics of individual workers or even the space of worker characteristics. Of course, in order for the firm to solve for its optimal (linear) payment schedule – that is, to solve for equilibrium of the two stage game – the firm must know more. We discuss this in examples.

In the general model, it does not seem possible to solve for stationary assortative equilibrium in closed form or to quantify the comparisons of firm profits in stationary assortative equilibrium and in the benchmarks. We therefore present examples in which solving in closed form and quantifying comparisons are possible; for these examples we can also determine the firm optimal linear payment schedule and determine how close how close stationary equilibrium comes to the profit the firm could achieve if it had full information and could tailor payment schedules to individual characteristics.[3]

Two features of our model deserve special attention. The first of this features is that the matching between workers and tasks is endogenous, and this endogeneity is a crucial driver of our conclusions. By contrast, in more familiar repeated game environments , either the matching is fixed – the same players interact in each period – or the matching is random. In these environments, the actions of players today affect the way in which others will play against them in the future, but not the future matching; in our setting,

[3]But solving for the optimal (possibly) non-linear payment schedule is beyond us, even in the most special of cases.

the actions of workers today affect how they will be matched in the future. The second of these features is that the dependence of output on effort, and hence the presence of moral hazard, has all the familiar implications – but it also has an important and unfamiliar implication: effort matters for optimal matching. To see why, suppose that all workers exerted the same effort. In that case, output would be supermodular in worker and task types and hence the optimal matching would be assortative, matching better workers to better tasks. But in our setting, all workers will not exert the same effort; the effort exerted by a particular worker will depend on the task to which it is matched, on the cost of effort and on the payment scheme. As a consequence, it may happen that when workers choose effort optimally the imputed output function will not be supermodular in worker and task types and hence the optimal matching will not assortative. indeed, as we show by example, the imputed output function may actually be *submodular* (rather than supermodular), so that the optimal matching will be *anti-assortative* (matching better workers to worse tasks), rather than assortative. In this case, the assortative matching will actually be *worst* not best. However, we show that, if the marginal cost of effort is log-concave, then the imputed output function will be supermodular and hence the optimal matching will indeed be assortative.

We are not aware of previous work on assignment models in which output depends on effort (in addition to the characteristics of workers and tasks) and so there seems no work that is very close to ours in economic terms. The work which seems closest to ours in mathematical terms is Hopkins (2012), which treats a static model in which output depends on the observable characteristics of firms and the unobservable characteristics of workers, so there is one-sided adverse selection. Hopkins' model unfold in two states: in the first stage, before matching takes place, workers have the opportunity to send costly signals which are observable; in the second stage, workers and firms are matched and production takes place. Hence the agents are playing a signaling game. Hopkins shows that there is a unique separating equilibrium of this signaling game and that this equilibrium leads to an assortative matching. His analysis relies on the results of Mailath (1987) to characterize signals in the separating equilibrium as the solutions to an ordinary differential equation that looks similar to ours.

Following this Introduction, Section 2 describes the environment in which we work. Section 4 describes our notion of stationary assortative equilibrium. Section 5 demonstrates existence and uniqueness of stationary assortative equilibrium, Section 6 presents natural benchmarks and qualitative comparisons with these benchmarks. Section 7 computes closed form solutions for a class of examples for which stationary assortative equilibrium can be computed in closed form and for which quantitative conclusions can be drawn with respect to the benchmarks. Section 8 offers a few concluding remarks. All proofs are collected in the Appendices.

2 Environment

We first introduce the basic framework of Tasks, Workers, Output, Payments and then collect the various assumptions.

There is a fixed space of *tasks* $T = [B, 1]$, where $B \geq 0$. Task t is characterized by its *quality* $q(t)$; we assume that $q : T \to [0, \infty)$ is smooth (by "smooth" we always mean twice continuously differentiable), (differentiably) strictly increasing and that $q(t) > 0$ if $t > 0$.[4] For convenience, we assume that the "population" of tasks is uniformly distributed and that the total mass of tasks is $1 - B$, this entails that the mass of tasks with quality less that that of task t is $(t - B)$. (The assumption that tasks are uniformly distributed according to their ranking entails little loss of loss of generality; the distribution of tasks by quality is implicitly defined by the function q.)

The space of *workers* is $W = [B, 1]$. Worker w is characterized by its *productivity* $p(w)$ and its *worker-specific cost factor* $k(w)$. We assume that $p : W \to [0, \infty)$ is smooth, (differentiably) strictly increasing and that $p(w) > 0$ if $w > 0$; we assume that $k : W \to [0, \infty)$ is smooth and weakly decreasing. For convenience, we assume that the population of workers is uniformly distributed; and that the total mass of workers is $1 - B$, so there are the same number of workers as tasks.[5] This entails that the mass of workers with productivity less than that of worker w is $(w - B)$. (Given

[4] In the examples, $B = 0$ and $q(0) = 0$; that is, the worst task is worthless.

[5] If there were more tasks than workers, the worst tasks would simply not be undertaken. If there were more workers than tasks, we would focus on an equilibrium in which the worst workers are unemployed, but the existence of unemployed workers would create additional complications.

the assumption that workers who are more productive also have lower cost factors, the assumption that workers are uniformly distributed according to their ranking again entails little loss of generality: the distribution of workers by productivity is implicitly defined by the functions p, k.)

If worker w is matched to task t and exerts effort $e \in E = [0, \infty)$ then it produces *output* $Y(e, w, t) = e\, p(w)q(t)$. (Given that output is multiplicatively separable, the assumption of linearity in effort is computationally convenient but innocuous. Workers choose effort e; if output were non-linear but strictly increasing and weakly concave in effort, so that $Y(e, w, t) = h(e)p(w)q(t)$, we could simply view workers as choosing *virtual effort* $h(e)$ rather than actual effort e. Linearity has the additional implication that, independent of the productivity of the worker and the quality of the task, aribtrarily large output can always be produced by exerting sufficient effort, but this fact plays no essential role.)

If worker $w \in W$ is matched to task $t \in T$, exerts effort $e \in [0, \infty)$ and produces output $Y = e\, p(w)q(t)$ it incurs the *cost* $C(e, w) = k(w)c(e)$ where $c : [0, \infty) \to [0, \infty)$ is a common *effort cost factor*. (Note that cost depends on effort and on worker type but not on task type. We could allow for cost that depends on effort, worker type and task type, provided that we made additional assumptions on the common cost factor $c(e)$.) We assume that c is smooth, strictly increasing and strictly convex; in order to guarantee that optimal choice of effort always exists (in particular, when workers behave myopically) we also assume that

$$\lim_{e \to 0} \frac{c(e)}{e} = 0 \quad \text{and} \quad \lim_{e \to \infty} \frac{c(e)}{e} = \infty$$

In view of the assumed convexity of c, this is equivalent to assuming that $c'(0) = 0$ and $\sup c'(e) = \infty$.[6]

If the worker produces output y then it receives a payment $P(y)$ where $P : [0, \infty) \to [0, \infty)$ is the *payment schedule*. By assumption, payment depends only on output. This seems natural: output is observable but effort and worker type are not. We require that $P(y) \leq y$ so that payment is a *share of output*. The realization of the payment is the *wage*, which will be determined endogenously in equilibrium.

For simplicity we focus here on linear payment schedules $P(y) = \lambda y$, with $\lambda \in (0, 1)$. This is largely for convenience/simplicity; much of the analysis would go through (with complications) for smooth weakly concave payment schedules. (Concavity is required in order that the optimization problem of workers have a unique solution.)

As usual, we assume utility is quasi-linear, so if in a given period, worker receives payment P and incurs cost C its net period utility is $P - C$.

Workers discount future utility at the constant rate $\delta \in (0, 1)$.[7] As usual, it is convenient to normalize, multiplying discounted lifetime rewards by $1-\delta$ and so expressing them in terms of discounted average per-period rewards. Hence, if in each period n, the worker exerts effort e_n and produces output

[6] If the domain of possible effort choices were bounded, the latter form of the assumptions would be more convenient.

[7] The assumption that workers share a common discount factor is familiar but is unnecessary. As we discuss at the end of Section 5, we could allow for worker-specific discount factors.

y_n, its lifetime utility will be

$$(1-\delta)\sum_{n=0}^{\infty}\delta^n[P(y_n)-C(e_n,w)]$$

2.1 Implications of the Assumptions

It may be helpful to note the important implications of these assumptions. For output Y the important implications are that for $e, w, t > 0$ we have:

- $Y > 0$
- Y is continuously differentiable and strictly increasing in each variable
- Y is strictly supermodular in each pair of variables
- Y is weakly concave in effort

For cost C the important implications are that

- C is continuously differentiable, strictly increasing and strictly convex in effort and weakly decreasing in worker type
- C is weakly submodular

For the interaction between output and cost the imporant implication is that

- optimal effort exists and is strictly positive

Finally, there is an implication for imputed effort. Write $\Phi(y, w, t) = y/p(w)q(t)$ for the effort required for worker w matched with task t to produce output y. (Note that Φ is well-defined and continuously differentiable when $w, t \neq 0$.) Direct calculation shows that

- the ratio $\frac{\partial\Phi/\partial y}{\partial\Phi/\partial t}$ is independent of w

In fact, these bulleted properties are really all that are necessary for much of our analysis, and so we could have assumed these properties directly, rather than assuming multiplicative separability. We have preferred to make the stronger assumptions because they make the (already complicated) proof considerably more transparent and less cumbersome and because the required property of the imputed effort function is hard to interpret and hard to verify except under the multiplicative separability assumptions we have made.

2.2 Information

It is common in game theoretic analysis to assume that the environment is common knowledge. In this setting, that would suggest that the quality $q(t)$ of all tasks, the productivity $p(w)$ and cost $k(w)$ of all workers and the common cost factor c – or at least the distribution of the objects – be common knowledge. However, at this point we do not require that agents have all this information. We require only that the firm know the *ordering* of tasks and observe the output of each worker (so that it can match workers to tasks), and that the workers know the quality of all tasks, their *own* productivity and cost, and the common cost factor. Indeed, the workers do not even need to know the spaces from which the productivity and cost functions of other workers are drawn.

3 Matching

In each period workers are matched with tasks, choose effort, produce output and receive payment. Because only output (not effort or worker type) is observed, it seems natural to assume that the matching of workers to task depends only on past history of output. We focus on the simplest and most obvious matching rule: workers are matched to tasks according to the ranking of output produced in the previous period. Recall that the spaces of workers and tasks are $W = [B, 1]$, $T = [B, 1]$ and that the total masses of workers and tasks are $1 - B$.

An *output mapping* is a (measurable) map $G : W \to [0, \infty)$; we interpret $G(w)$ as the output produced by worker $w \in W$. An *output distribution* is a (measurable) mapping $\Gamma : [0, \infty) \to [0, 1 - B]$; we interpret $\Gamma(y)$ as the mass of workers who produce output at most y (so Γ is a cumulative distribution function). Given an output mapping G, the corresponding output distribution Γ is defined by setting $\Gamma(y)$ to be the (Lebesgue) measure of the set $\{w \in W : G(w) \leq y\}$. Notice that the output distribution Γ is unaffected by changing G on a set of measure 0 and in particular is independent of the output of any single worker.

Fix the current output distribution Γ and an output $y \in [0, \infty)$. If worker w produces output y in the current period then in the next period the *matching rule* μ assigns worker w to task $t = \mu(y) = B + \Gamma(y)$. That is, worker w is assigned to the task whose rank in the task distribution is precisely the same as the rank of y in the output distribution. In particular, if y is the worst output, then w will be assigned the worst task, and so forth.

This matching rule requires some comment. Consider an output map G and the corresponding output distribution Γ. As we have already noted, the distribution Γ is independent of the output of any single worker, so the output choice of worker w affects the task assigned to worker w but not the task assigned to any other worker. (Of course this is one reason we have chosen to work in a continuum model.) If G is not one-to-one then there will be workers $w, \hat{w}$ for which $G(w) = G(\hat{w})$, so workers $w, \hat{w}$ will be assigned the same task. However because we focus on equilibria in which the output mapping G *will be* one-to-one, we can ignore this complication. Even if the actual output mapping G is one-to-one, it may be that worker w contemplates a deviation from G in which he/she produces output $y \neq G(w)$; in particular, worker w may contemplate producing the output $G(\hat{w})$ of some other worker $\hat{w}$. In that case, the matching rule would again assign workers $w, \hat{w}$ to the same task. This seems an unavoidable complication of the continuum model. (In a large finite model, worker w could simply produce an output slightly greater than $G(\hat{w})$ and this problem would not arise – but then worker w's choice would affect the task assigned to worker $\hat{w}$.) Because this will occur only as a counter-factual, we will ignore this potential complication as well.

4 Stationary Assortative Equilibrium

In each period, each worker w is assigned a task and chooses an effort level to exert. At the end of each period the entire output distribution is revealed.[8] The history of worker w in period n is therefore the sequence of previous assignments of tasks, choces of effort and observed output distributions. In the current period, the worker is assigned a task and must choose an effort level, so a *(pure) strategy* for worker w is therefore a map $\sigma_w : \textit{history} \times T \to E$ from (past) history and (current) task to effort. Note that, given worker type

[8] Note that workers observe the output distribution but not necessarily the output mapping. In particular workers know how many other workers produced output below a given level – but not the names of those workers.

w and task type t, effort e determines output $y = Y(e, w, t) = ep(w)q(t)$ and output determines effort $e = \Phi(y, w, t) = y/p(w)q(t)$ so so there is no loss in viewing a strategy as a map from from history and task to *output*.

Now fix a strictly increasing output mapping G and the corresponding output distribution Γ. Fix a worker w and a strategy σ_w for worker w. Suppose that other workers produce according to G in every period and in particular that the output distribution is Γ in every period. If worker w is initially assigned to the task t_0 in period 0, then the strategy σ_w determines the output y_0 to be produced in period 0, and the effort $e_0 = y_0/p(w)q(t_0)$ required to produce this output. The output level y_0 will place worker w at some point in the output distribution Γ and hence determine the task $t_1 = \mu(y_0)$ to which worker w will be assigned in period 1. The strategy σ_w determines the output y_1 to be produced in period 1 and the effort $e_1 = y_1/p(w)q(t_1)$ required, and so forth. Thus, assuming that other workers produce in such a way that the output distribution is Γ in every period, the strategy σ_w and the initial task assignment t_0 determines the entire history of worker w: in period n, worker w is assigned task t_n and produces output y_n which requires effort $e_n = y_n/p(w)q(t_n)$. This yields worker w the utility $\lambda y_n - k(w)c(e_n)$ in period n and hence lifetime utility of

$$V_w(\sigma|t_0, \Gamma) = (1-\delta)\sum_{n=0}^{\infty} \delta^n [\lambda y_n - k(w)c(e_n)]$$

The strategy σ_w is *optimal from* t_0 (for worker w) if

$$V_w(\sigma_w|t_0, \Gamma) = \sup V_w(\widehat{\sigma}_w|t_0, \Gamma)$$

where the supremum is taken over all pure strategies $\widehat{\sigma}_w$.

A *stationary assortative equilbrium* consists of a strictly increasing output map G and strategies σ_w for each worker such that

- $\sigma_w(t, h) = \Gamma(w)$ for every task t and history h
- σ_w is optimal from $t_0 = w$

That is: workers produce output as prescribed by G in every period and find it optimal to do so. Given this behavior of workers, the matching rule guarantees that, along the equilibrium path, matching is assortative in every period (worker w is matched with task $t = w$). As usual, we require optimality among *all* strategies, not just among constant or stationary strategies, so we allow for the possibility that worker w contemplates a strategy that calls for a complicated sequence of production plans , which would lead to a correspondingly complicated sequence of assignments to tasks. Because behavior in a stationary assortative equilibrium is completely determined by the output mapping G, we identify the equilibrium with G itself. This should cause no confusion.

5 Existence and Uniqueness of SAE

Our fundamental result is that, if $B > 0$ (so that even the worst worker when matched to the worst task can produce strictly positive output) then there exists a unique stationary assortative equilibrium.

Theorem 1 *Fix the payment schedule* $\lambda \in (0, 1)$*. If* $B > 0$ *(so that even the worst worker matched to the worst task can produce strictly positive output) then:*

(i) there is a unique stationary assortative equilibrium

(ii) the equilibrium output mapping G is continuously differentiable and is the unique increasing solution to the ordinary differential equation

$$G'(w) = \delta \left(\frac{k(w)q'(w)c'\left(\frac{G(w)}{p(w)q(w)}\right)}{q(w)\left[k(w)c'\left(\frac{G(w)}{p(w)q(w)}\right) - \lambda p(w)c(w)\right]} \right) G(w) \qquad (1)$$

subject to the initial condition that the worst worker, matched to the worst task, chooses the level of output that maximizes current utility

(iii) worker utility is strictly increasing in worker type w.

5.1 Comments

We have assumed in that all workers discount the future at the same rate δ, but this assumption is not necessary. We could allow worker-specific discount factors $\delta(w)$, provided that the map $\delta : W \to (0,1)$ is smooth and satisfies natural assumptions; the only change would be that in the ordinary differential equation (1) the factor δ would become a factor $\delta(w)$.

The assumption that $B > 0$ means that the worst worker and worst task are not worthless, which seems natural. However, it is also of interest to treat the setting in which the worst worker and worst task are worthless, not least because – as we show in Section 7 – it is only in that setting that we can find solutions in closed form (for specific functional forms). Unfortunately, in the latter setting, the ODE for the solution and for the inverse of the solution are no longer Lipschitz, so existence and uniqueness of solutions are not guaranteed. In fact, existence can be proved by a limiting argument, but uniqueness requires an additional assumption; see Section 7.

As we have noted earlier, we do not require that workers know very much about the environment, and hence there is no reason to suppose that workers can solve for stationary assortative equilibrium. Equilibrium simply describes a particular state of the system; our knowledge of the parameters allows us to solve for this particular state of the system, but we do not offer any mechanism by which the system reaches this state.

6 Profit Benchmarks

It seems natural to suppose that the firm seeks to maximize (expected) profit: the (expected) output produced minus (expected) payments. For a given fixed payment schedule $P(y) = \lambda y$, we compare profit Π_{SAE} in the stationary assortative equilibrium G against three natural benchmarks.

- Π_{random} is what the firm's (expected) profit would be if the firm committed to matching workers and tasks *randomly* in each period and paying according to the payment schedule $P(y) = \lambda y$, and workers then chose effort optimally given these commitments.
- Π_{assort} is what the firm's profit would be if the firm could actually observe workers' characteristics, committed to matching workers and tasks *assortatively* in each period and paying according to the payment schedule $P(y) = \lambda y$, and workers then chose effort optimally given these commitments.
- Π_{FI} is what the firm's profit would be if the the firm could actually observe workers' characteristics and could use this information to match

workers and tasks and to set a worker/task-specific payment schedule, and workers then chose effort optimally, given these assignments and payment schedules.

It must be kept in mind that, in our setting, the firm cannot observe workers' characteristics so only the first benchmark represents something that the firm could actually achieve. However comparing firm profit in the stationary assortative equilibrium with these benchmarks provides a useful measure of how much the firm is able to achieve with the limited power given to it in comparison with what it might achieve if it had greater power.

It is not hard to see that $\Pi_{\text{assort}} \leq \Pi_{\text{SAE}}$ and that $\Pi_{\text{SAE}} \leq \Pi_{\text{FI}}$. (In fact these inequalities are necessarily strict.) Because output is assumed to be supermodular, it might seem obvious that $\Pi_{\text{random}} \leq \Pi_{\text{assort}}$ – but in fact this need not be so. The reason, to which we have already alluded in the Introduction, is that, although output $Y(e, w, t)$ is supermodular (in every pair of variables), when worker w is matched to task t and chooses effort optimally , the *imputed output* $Y^*(w, t)$ might not be supermodular in w, t – in which case the optimal matching (i.e. the matching that maximizes total output) will *not* be assortative. Indeed, for some functional forms for output Y and cost C and some payment schedules $P() = \lambda y$, imputed output $Y^*(w, t)$ will be *submodular* in w, t; in that case *anti-assortative matching will be the optimal matching* and *assortative matching will be the worst matching*; in particular, assortative matching will yield lower profit than random matching. (We defer the example and calculations to the Appendix C.)

However, if marginal cost is (weakly) log-concave in effort – i.e., $\log(\partial C/\partial e)$ is (weakly) concave with respect to e – then imputed output $Y^*(w, t)$ can be shown to be supermodular in w, t and assortative matching will be optimal.[9] This leads to the profit comparisons that intuition suggests.

Theorem 2 *If marginal cost is log-concave in effort, then firm profits in the stationary assortative equilibrium and the benchmarks are ordered as follows:*

$$\Pi_{\text{random}} < \Pi_{\text{assort}} < \Pi_{\text{SAE}} < \Pi_{FI}$$

(Note that the inequalities are strict.)

The comparisons of $\Pi_{\text{random}}, \Pi_{\text{assort}}, \Pi_{\text{SAE}}$ in Theorem 2 take the payment rule $P(y) = \lambda y$ as fixed. If the firm is a monopolist, it would seem natural to assume that it should choose the payment rule in order to maximize profit Π_{SAE}. However, if the firm does not know worker's production functions or cost functions – or even the distributions of these functions – it does not seem clear how the firm should go about choosing a payment rule to maximize profit – or indeed, even what it means for the firm to maximize profit. We ~~return to this point in the next~~ Section.

[9]Because cost $C(e, w) = c(e)k(w)$ is separable, marginal cost is (weakly) log-concave in effort exactly when $c'(e)$ is log-concave; if c is three-times continuously differentiable this reduces to:

$$[\log c'(e)]'' = \frac{c'(e)c'''(e) - c''(e)^2}{c'(e)^2} \leq 0$$

Because we have already assumed $c'(e) > 0$ this is equivalent to the assumption that $c'(e)c'''(e) \leq c''(e)^2$ which in turn is equivalent to the assumption that $\frac{c''(e)}{c'(e)}$ is weakly decreasing. Note that this relatively weak assumption is satisfied by the cost functions $c(e) = e^{\alpha}$ (for $\alpha > 1$) which we will consider in when we compute examples in the next section.

7 Examples

To illustrate the general results of Theorems 1 and 2, it seems useful to consider examples for which solutions can be computed in closed form. Unfortunately, even for the simplest functional forms it seems impossible to solve the differential equation (1) in closed form when $B > 0$ (so that the initial condition is that the worst worker is matched to the worst form and exerts the myopically optimal effort). To get around this, problem, we take $B = 0$ and consider a class of functional forms for which the worst worker and the worst firm are worthless (i.e., no level of effort can produce positive output). In this setting, we can explicitly write down closed form solutions, prove that the solutions are unique (a fact that no longer follows from standard uniqueness results) and show that the solution when $B = 0$ approximates the solution when $B > 0$ but small.

For the remainder of this Section, we assume $B = 0$ so the worker space is $W_0 = [0, 1]$ and the task space is $T_0 = [0, 1]$. We begin with the simplest functional forms: $Y(e, w, t) = ewt$, $C(e, w) = e^2$, $P(y) = \lambda y$ for $\lambda \in (0, 1)$; the analysis for more general functional forms (discussed below) is almost the same although the algebra is much messier. Note that the worst worker and the worst task are worthless. We assert that, for this setting and these functional forms, there is a unique stationary assortative equilibrium:

$$G_0(w) = \left(\frac{2\lambda}{4 - \delta} \right) w^4$$

Moreover, if for each $B > 0$ we denote by B_B the unique stationary assortative equilibrium guaranteed by Theorem 1 when we keep output, cost, payment rule the same but restrict the worker and task spaces to $W_B =$

$[B,1], T_B = [B,1]$ then $G_B \to G_0$ and $G'_B \to G'_0$ uniformly on every interval $[b,1]$ with $b > 0$.

To see that G_0 is the unique stationary assortative equilibrium, note first that the same argument as in the proof of Theorem 1 shows that every stationary assortative equilibrium G is smooth, strictly increasing and satisfies the ODE (1), and that every strictly increasing solution to the ODE (1) is a stationary assortative equilibrium. For the given functional forms, the ODE (1) reduces to

$$G'(w) = \frac{2\delta G(w)^2}{[2G(w) - \lambda w^4]\, w} \tag{2}$$

Direct computation shows that the function G_0 solves (2) and satisfies the initial condition $G_0(0) = 0$ so to show that G_0 is the unique stationary assortative equilibrium it remains only to show that it is the unique increasing solution to (2) satisfying the initial condition. To see this, suppose that $\widehat{G}_0$ were another solution satisfying the initial condition. The ODE (2) is Lipschitz away from the critical curve where the denominator $[2G(w) - \lambda w^4]\, w$ of the right hand side of (2) is zero, so the solutions $G_0, \widehat{G}_0$ cannot cross for $w \neq 0$. In particular, if $G_0(w) > \widehat{G}_0(w)$ for some w then $G_0(w) > \widehat{G}_0(w)$ for all w, and vice versa. However, it is easily checked that the right hand side of (2) is strictly decreasing in G, so if $G_0(w) > \widehat{G}_0(w)$ for all w it would necessarily be the case that $G'_0(w) < \widehat{G}'_0(w)$ for all w and vice versa, which would violate the Mean Value Theorem. Hence G_0 is the unique unique increasing solution to (2) satisfying the initial condition and hence the unique stationary assortative equilibrium.

To see that $G_B \to G_0$ and $G'_B \to G'_0$, note that, because solutions to the ODE (2) cannot cross it must be the case that the solutions G_B are strictly decreasing in B (i.e. $G_B(w) < G_{\hat{B}}(w)$ if $B > \hat{B}$ and $w \geq B$) and hence converge to *some* function F. The fact that the solutions G_B solve the ODE (2) guarantee that, away from the critical curve, the solutions G_B are equi-uniformly differentiable (i.e., the difference quotients converge to the derivative at a rate that is independent of $w \in [b, 1]$ provided that $0 < b < B$), and hence that the the functions G_B and their derivatives G'_B converge uniformly to F and F' (respectively) for $w \in [b, 1]$. It follows that the limit function F satisfies the ODE (2) on $(0, 1]$ and that $\lim_{w \to 0} F(w) = 0$, and hence that $F = G_0$, so we obtain the desired convergence assertion.

We can now compute the firm profit for the stationary assortative equilibrium and the benchmarks. Fix $\lambda \in (0, 1)$.

- If workers and tasks are always matched randomly and the payment schedule $P(y) = \lambda y$ is fixed, then worker w matched with task t will choose effort to maximize $\lambda ewt - e^2$ and hence will choose effort $e = \lambda wt/2$ and produce output $= \lambda w^2 t^2/2$. Because the firm retains the fraction $1 - \lambda$ of output, firm profit will be

$$\Pi_{\text{random}} = \int_0^1 \int_0^1 (1-\lambda)(\lambda w^2 t^2/2)\, dt\, dw = [(1-\lambda)\lambda][1/18]$$

- If workers and tasks are always matched assortatively and the payment schedule $P(y) = \lambda y$ is fixed, then worker w matched with task w will choose effort to maximize $\lambda e w^2 - e^2$ and hence will choose effort $e = \lambda w^2/2$ and produce output $= \lambda w^4/2$. Firm profit will be

$$\Pi_{\text{assort}} = \int_0^1 (1-\lambda)[\lambda w^4/2]\, dw = [(1-\lambda)\lambda][1/10]$$

- In the stationary assortative equilibrium, worker w is matched with task $t = w$ and produces output $G_0(w) = [(2\lambda)/(4-\delta)]\, w^4$ so firm profit will be

$$\Pi_{\text{SAE}} = \int_0^1 (1-\lambda)(2\lambda/(4-\delta)]w^4)\, dw = [(1-\lambda)\lambda][2/5(4-\delta)]$$

- Finally, we consider the full information firm optimum; i.e. the profit the firm would make if it knew the characteristics of each worker and could offer worker/task-specific payment schedules. In that case, the firm can extract the full surplus from each worker. If worker w were matched to task t the firm extracts the full surplus by offering a payment schedule to maximize profit (output net of payment) subject to the incentive constraint on worker effort. The firm would therefore induce the effort level e that maximizes $ewt - e^2$. This effort level is $e = wt/2$ so optimal profit would be $\pi(w,t) = w^2t^2/2 - w^2t^2/4 = w^2t^2/4$. Note that the function $\pi(w,t)$ is supermodular in w, t so the matching that yields optimal profit is assortative. Hence in the full-information firm optimum, worker w is matched to task $t = w$, produces output $w^4/2$ and receives wage $w^4/4$. Firm profit in the full-information optimum is

$$\Pi_{\text{FI}} = \int_0^1 \left(w^4/4\right)\, dw = 1/20$$

Because $\delta \in (0,1)$ we see that $1/10 < 2/5(4-\delta) < 2/15$; because $\lambda \in (0,1)$, we see that $(1-\lambda)\lambda \leq 1/4$ so

$$\Pi_{\text{random}} < \Pi_{\text{random}} < \Pi_{\text{SAE}} < \Pi_{\text{FI}}$$

Evidently, $\Pi_{\text{random}}, \Pi_{\text{random}}, \Pi_{\text{SAE}}$ are all maximized by taking $\lambda = 1/2$; noting that $\lim_{\delta \to 1} 2/5(4-\delta) = 2/15$ we see that if the firm chooses the optimal (linear) payment schedule and workers are perfectly patient then we have

$$\begin{array}{ccccccc}
\Pi_{\text{random}} & < & \Pi_{\text{random}} & < & \Pi_{\text{SAE}} & < & \Pi_{\text{FI}} \\
\| & & \| & & \| & & \| \\
\frac{1}{72} & < & \frac{1}{40} & < & \frac{1}{30} & < & \frac{1}{20}
\end{array}$$

A similar analysis can be carried through for the wide class of functional forms $Y(e,w,t) = ew^a t^b$, $C(e,w) = e^d w^{-s}$, $P(y) = \lambda y$ (assuming $a, b > 0$, $a + b \geq 1$, $d \geq 2$, $s \geq 0$). The unique stationary assortative equilibrium has the form $G_0(w) = Aw^\gamma$, where

$$\begin{aligned}
\gamma &= \frac{(a+b)d - s}{d-1} \\
A &= \left(\frac{\lambda\gamma}{d\gamma - \delta b} \right)^{1/(d-1)}
\end{aligned}$$

Solving for the profit in the first two benchmarks and in the stationary assortative equilibrium shows that each of Π_{random}, Π_{assort}, Π_{SAE} is the product of $(1-\lambda)\lambda^d$ and terms that involve only the exponents a, b, d, s and the discount factor δ but do not involve λ. (We leave the straightforward but messy algebra to the reader.) Hence for each of these benchmarks, the optimal (linear) payment schedule maximizes $(1-\lambda)\lambda^d$, which is accomplished by taking $\lambda = d/(1+d)$. Thus the optimal (linear) payment schedule depends only on the common cost factor $c(e) = e^d$ and not on the the productivity $p(w) = w^a$ of workers, the quality of tasks $q(t) = t^b$ or the worker-specific cost factor $k(w) = w^{-s}$. It is interesting to speculate to what extent this conclusion depends on the specific functional forms (or on the homogeneity of c) and to what extent it remains true more generally.

7.1 Comment

We noted in Subsection 5.1 that existence and uniqueness of solutions to the ODE – and hence existence and uniqueness of SAE – presents additional complictions when the worst worker and worst task are worthless. The examples above address these complications by demonstrating that solutions G_B for the restricted domains $W_B = [B, 1], T_B = [B, 1]$ converge to solutions G_0 on the domains $W_0 = [0, 1], T_0 = [0, 1]$. This argument is perfectly general, so that existence of SAE is guaranteed even when the worst worker and worst task are worthless. Uniqueness seems more complicated however. In the examples, uniqueness is derived from the fact that the right-hand side of the ODE is decreasing in G. Unfortunately, this fact does not hold for more general functional forms. A sufficient condition is that the common cost factor satisfies $c'(e) \leq c(e)/e + c(e)c''(e)/c'(e)$. The reader can easily verify that this inequality is satisfied when $c(e) = e^{\alpha}$ for $\alpha > 1$, as in the examples, but it does not follow from our other assumptions and does not seem to have any obvious intuitive interpretation.

8 Conclusion

In this paper we have formulated and analyzed a dynamic assignment model with one-sided adverse selection (unobserved worker characteristics) and moral hazard (unobserved worker effort). For this environment we have defined a notion of stationary equilibrium in which workers are matched to tasks endogenously on the basis of observable output and shown that (for a given payment schedule) such an equilibrium exists and is unique. At equilibrium, adverse selection is eliminated and moral hazard is mitigated. Firm profit in equilibrium is compared with natural benchmarks. For specific examples, we find closed-form solutions and solve for the optimal (linear) payment scheme.

In the environment we have considered here, a single firm owns a large family of tasks and outsources them to a large family of workers in each period. In this environment, we have assumed that the single firm sets a payment schedule that depends only on observed output and the firm matches tasks to workers. The equilibrium matches workers to tasks and specifies output (equivalently effort) for each worker (in each period). Equilibrium is driven by competition among workers.

A natural extension of this environment would consider a family of firms each of which owns a single task which it seeks to have performed by a single worker each period. In this extension, it would be natural to assume that each firm sets a payment schedule that depends on observed output but is specific to the particular firm/task, and that some central agency/platform matches tasks to workers. In this environment, it seems natural to insist that the payment schedules set by firms be determined as part of the equilibrium, and determined in equilibrium by competition among firms. Gretsky, Ostroy and Zame (1999) provides some hints as to how this might occur.

Another extension that seems natural would consider not workers and tasks but workers of different kinds with complementary skills, so that the issue is matching complementary workers in teams. In that environment, it seems natural to contemplate adverse selection and moral hazard on both sides, which would make analysis very challenging indeed.

Appendix A: Proof of Theorem 1

The proof of Theorem 1 is in two parts. In the first part we introduce the (weak) notion of imitative equilibrium and use results of Mailath (XXXX) to show that there is a unique imitative equilibrium, that the imitative equilibrium satisfies the ODE (1), and that in the imitative equilibrium, worker utility is strictly increasing in worker index. In the second – and much more complicated – part, we show that the unique imitative equilibrium is in fact a stationary assortative equilibrium.

To define an imitative equilibrium, fix a strictly increasing output function G that satisfies the initial condition that the worst worker, matched with the worst task, produces the myopically optimal output. For each worker w, task τ and output level y define $V(w, \tau, y)$ to be the long-run utility of worker w when it is matched to task w in period 0 and to task τ in every succeeding period and produces output y in every period. Keeping in mind that the effort required to produce output y depends on the task, we see that

$$\begin{aligned} V(w,\tau,y) &= (1-\delta)\left[\lambda y - C\left(\Phi(y,w,w),w\right)\right] + \delta\left[\lambda y - C\left(\Phi(y,w,\tau),w\right)\right] \\ &= (1-\delta)\left[\lambda y - k(w)\, c\left(\frac{y}{p(w)q(w)}\right)\right] \\ &\quad + \delta\left[\lambda y - k(w) c\left(\frac{y}{p(w)q(\tau)}\right)\right] \end{aligned}$$

If $y \in G([B,1])$ then $y = G(\hat{w})$ is the output specified for worker $\hat{w} = G^{-1}(y)$. Because G is strictly increasing, the matching rule matches a worker producing output $G(\hat{w})$ to the task $\hat{w}$. If worker $w$produces output y in every period then in period 1 and in every succeeding period worker w will be matched to task $\hat{w} = G^{-1}(y)$. Thus, in following this strategy, worker w is imitating worker $\hat{w}$.

We say G is an *imitative equilibrium* if no worker can gain by imitating another worker. That is, for every $w \in W$ we have the equivalent conditions

$$\begin{aligned} G(w) &\in \operatorname{argmax}_{y \in G([B,1])} V\left(w, G^{-1}(y), y\right) \\ w &\in \operatorname{argmax}_{\hat{w} \in W} V\left(w, \hat{w}, G(\hat{w})\right) \end{aligned}$$

Our first task is to prove that a unique imitative equilibrium equilibrium exists and establish some of its properties. The key is that when we restrict to imitation strategies, we have turned the original infinite horizon game into a signaling game to which the results of Mailath can be applied once we verify the necessary properties of the value function V. We formalize all of this as a Lemma.

Lemma 1 *There is a unique imitative equilibrium G. It is strictly increasing, continuously differentiable, and satisfies the ODE 1 with the specified initial condition, and worker utility is strictly increasing in worker index.*

Proof We first show that the function V satisfies Mailath's conditions.

Condition (1): "smoothness" Based on our assumptions, the functions P, C, Φ are all twice continuously differentiable. As a result, $V(w, \tau, y)$ is twice differentiable on $[B, 1]^2 \times \mathbb{R}$.

Condition (2): "belief monotonicity" Differentiating V with respect to τ yields

$$\frac{\partial V(w, \tau, y)}{\partial \tau} = \delta \left[\frac{y k(w) q'(\tau)}{p(w) \left[q(\tau)\right]^2} \right] c' \left(\frac{y}{p(w) q(\tau)} \right) > 0$$

for all $w \in W$, $\tau \in T$ and $y > 0$. This is Condition (2).

Condition (3): "type monotonicity" Differentiating V with respect to τ and then with respect to y yields

$$\begin{aligned}\frac{\partial^2 V(w,\tau,y)}{\partial w \partial y} &= (1-\delta)\left[\frac{-k'(w)}{p(w)q(w)} + \frac{k(w)p'(w)}{[p(w)]^2 q(w)} + \frac{k(w)q'(w)}{p(w)\,[q(w)]^2}\right] c'\left(\frac{y}{p(w)q(w)}\right) \\ &+ (1-\delta)\left(\frac{yk(w)}{[p(w)q(w)]^3}\right)[p'(w)q(w) + p(w)q'(w)]\, c''\left(\frac{y}{p(w)q(\tau)}\right) \\ &+ \delta\left[\frac{-k'(w)p(w) + k(w)p'(w)}{[p(w)]^2 q(\tau)}\right] c'\left(\frac{y}{p(w)q(\tau)}\right) \\ &+ \delta\left[\frac{yk(w)p'(w)}{[p(w)q(\tau)]^3}\right] c''\left(\frac{y}{p(w)q(\tau)}\right)\end{aligned}$$

Keeping in mind that k is weakly decreasing, so that $k' \leq 0$, we see that each of the terms on the right-hand side are positive, so $\frac{\partial^2 V(w,\tau,y)}{\partial w \partial y} > 0$ for all $w \in W$, $\tau \in T$ and $y > 0$. This is Condition (3).

Condition (4) Recalling that the function c is strictly convex, we have

$$\frac{\partial^2 V(w,w,y)}{\partial y^2} = -\left(\frac{k(w)}{[p(w)q(w)]^2}\right) c''\left(\frac{y}{p(w)q(w)}\right) < 0$$

for all $w \in W$, $\tau \in T$ and $y > 0$. Thus, $V(w,w,y)$ is strictly concave in y.

Moreover, we have

$$\frac{\partial V(w,w,y)}{\partial y} = \lambda - \left[\frac{k(w)}{p(w)q(w)}\right] c'\left(\frac{y}{p(w)q(w)}\right)$$

The right-hand side is strictly positive at $y = 0$ and is strictly decreasing in y, so for each w the equation $\frac{\partial V(w,w,y)}{\partial y} = 0$ has a unique solution in y, and the unique solution maximizes $V(w,w,y)$. Together, these are Condition (4).

Condition (5): "boundedness" We have already noted that $V(w,w,y)$ is strictly concave in y so there does not exist any $w \in W$ and $y \geq 0$ such that $\frac{\partial^2 V(w,w,y)}{\partial y^2} \geq 0$. Hence, Condition (5) is satisfied.

Since $V(w, \tau, y)$ is increasing in τ, the initial condition in [**?**] (i.e., Condition (6)) is that $G(B) = \arg\max_y V(B, B, y)$, namely the worst worker B, when matched to the worst task B, chooses myopically optimal output.

We have verified Mailath's Conditions (1)-(5) and we are restricting to an output function G that satisfies the initial condition so Mailath's Theorems 1 and 2 guarantee that there is every imitative equilibrium G is continuous on $[B, 1]$, smooth and strictly monotonic on $(B, 1)$ and is solves the ODE

$$G'(w) = -\frac{\left.\frac{\partial G(w,\tau,y)}{\partial \tau}\right|_{\tau=w}}{\frac{\partial G(w,w,y)}{\partial y}}$$

Straightforward calculation shows that this ODE coincides with (1).

Mailath's Theorem 2 shows that $G'(w)$ has the same sign as $\frac{\partial V(w,\tau,y)}{\partial w \partial y}$, which we have already shown to be positive; hence every imitative equilibrium is strictly increasing. Moreover, because

$$\left.\frac{\partial V(w, \tau, y)}{\partial \tau}\right|_{\tau=w} = \delta \left[\frac{k(w)q'(w)}{p(w)\left[q(w)\right]^2} y\right]$$

which is bounded for each $w \in W$, $y \in [0, \infty)$ so the ODE has a unique solution that satisfies the initial condition.

We can therefore conclude that there is a unique imitative equilibrium G, that G is smooth and strictly increasing, and that G is the unique solution to the ODE 1 with the initial condition that $G(B)$ is worker B's myopically optimal output when matched to task B.

Finally, to show that worker utility is increasing in worker type note that in an imitative equilibrium, worker w receives the same payoff in each period so its long-run utility is $U(w) = \lambda G(w) - k(w)c\left(\frac{G(w)}{p(w)q(w)}\right)$. Differentiating and doing the requisite algebra yields

$$
\begin{aligned}
U'(w) &= \lambda G'(w) - k'(w)c\left(\frac{G(w)}{p(w)q(w)}\right) \\
&\quad - k(w)c'\left(\frac{G(w)}{p(w)q(w)}\right)\left[\frac{G'(w)}{p(w)q(w)} - \frac{G(w)p'(w)}{[p(w)]^2\, q(w)} - \frac{G(w)q'(w)}{p(w)\,[q(w)]^2}\right] \\
&= \left[\lambda - \frac{k(w)c'\left(\frac{G(w)}{p(w)q(w)}\right)}{p(w)q(w)}\right] G'(w) - k'(w)c\left(\frac{G(w)}{p(w)q(w)}\right) \\
&\quad + k(w)c'\left(\frac{G(w)}{p(w)q(w)}\right) G(w)\left[\frac{p'(w)}{[p(w)]^2\, q(w)} + \frac{q'(w)}{p(w)\,[q(w)]^2}\right] \\
&= -\delta k(w)c'\left(\frac{G(w)}{p(w)q(w)}\right) G(w)\frac{q'(w)}{p(w)\,[q(w)]^2} \\
&\quad - k'(w)c\left(\frac{G(w)}{p(w)q(w)}\right) \\
&\quad + k(w)c'\left(\frac{G(w)}{p(w)q(w)}\right) G(w)\left[\frac{p'(w)}{[p(w)]^2\, q(w)} + \frac{q'(w)}{p(w)\,[q(w)]^2}\right] \\
&= k(w)c'\left(\frac{G(w)}{p(w)q(w)}\right) G(w)\left[\frac{p'(w)}{[p(w)]^2\, q(w)} + \frac{(1-\delta)q'(w)}{p(w)\,[q(w)]^2}\right] \\
&\quad - k'(w)c\left(\frac{G(w)}{p(w)q(w)}\right)
\end{aligned}
$$

Since k is weakly decreasing, this last expression is strictly positive, so worker utility is increasing in worker type w. This completes the proof. ■

Lemma 1 constitutes the first part of the proof of Theorem 1. We now turn to the second part, showing that an imitative equilibrium is a stationary assortative equilibrium. We arrange the proof as a sequence of lemmas.

By definition, an imitative equilibrium has the property that no worker has a profitable deviation that consists of imitating some other worker; we must show that no worker has any profitable deviation at all. We first show that if a profitable deviation exists then there is a profitable finite deviation.

This is a simple and familiar consequence of discounting so we omit the proof.

Lemma 2 *Let G be the unique imitative equilibrium. If worker w has a profitable deviation from G then worker w has a profitable deviation from G in which, after a finite number of periods, it produces output $G(w)$ forever.*

We now study optimal finite-period deviations. Because G is continuous, the range of G is precisely the interval $[G(B), G(1)]$. First, we show that it is dominated for workers to choose output outside the interval $[G(B), G(1)]$ in any period.

Lemma 3 *Choosing output $y \notin [G(B), G(1)]$ is dominated in every period.*

Proof First consider the case in which the worker contemplates producing output less than $G(B)$. In view of the initial condition, $G(B)$ is worker B's myopically optimal output when matched to task B. We have shown that $G(B)$ is the solution to the equation

$$\frac{\partial V(B,B,y)}{\partial y} = 0.$$

Hence

$$\lambda - \left[\frac{k(B)}{p(B)q(B)}\right]\left[c'\left(\frac{G(B)}{p(B)q(B)}\right)\right] = 0.$$

Moreover, for every $w \in [B,1]$, $\tau \in [B,1]$, and $y < G(B)$ we have

$$\begin{aligned}
\frac{\partial\left[\lambda y - c\left[y/p(w)q(\tau), w\right) k(w)\right]}{\partial y} &= \lambda - \frac{k(w)}{p(w)q(\tau)}c'\left(\frac{y}{p(w)q(\tau)}\right) \\
&> \lambda - \frac{k(B)}{p(B)q(B)}c'\left(\frac{G(B)}{p(B)q(B)}\right) \\
&= 0,
\end{aligned}$$

That is, worker w's current payoff is strictly increasing in its output at $y < G(B)$. Hence, if in period T, worker w produces output $y < G(B)$ then its current payoff will be less than if it produced output $G(B)$. Moreover, since it will be producing the worst output of any worker, it will be matched in period $T+1$ to the worst task (i.e. task B) – exactly as if it had produced output $G(B)$. Hence producing output $y < G(B)$ in period T is dominated by producing output $G(B)$.

Now consider the case in which the worker contemplates producing output greater than $G(1)$. Since $G(w)$ is strictly increasing, and since the numerator of the ODE 1 is positive, the denominator of the ODE must also be positive, which is equivalent to

$$k(w)c'\left(\frac{G(w)}{p(w)q(w)}\right) > \lambda p(w)q(w),\ \forall w \in [B,1].$$

If worker $w \in [B,1]$ is matched to task $\tau \in [B,1]$ then for every $y > G(1)$, we have

$$\begin{aligned}\frac{\partial\left[\lambda y - c\left(y/p(w)q(\tau), w\right)k(w)\right]}{\partial y} &= \lambda - \frac{k(w)}{p(w)q(\tau)}c'\left(\frac{y}{p(w)q(\tau)}\right) \\ &= \lambda - \frac{k(1)}{p(1)q(1)}c'\left(\frac{G(1)}{p(1)q(1)}\right)\end{aligned}$$

The last expression is negative, so we see that worker w's current payoff is strictly decreasing in its output at $y > G(1)$. Hence if in period T worker w produces output $y > G(1)$ then its current payoff will be less than if it produced output $G(1)$. Moreover, since it will be producing the best output of any worker, it will be matched in period $T+1$ to the best task (i.e. task a) – exactly as if it had produced output $G(a)$. Hence producing output $y > G(1)$ in period T is dominated by producing output $G(1)$.[10] ■

[10]It is worth noting that there is nothing special about the output $G(1)$ in the sense

In view of the preceding lemmas, we can focus on finite deviations in which the worker w produces output in $[G(B), G(1)]$ in every period and returns to producing output $G(w)$ from some point on. Each such finite deviation is characterized by a sequence of periods $0, 1, \ldots, N$ in which the worker is matched with tasks $w_0 = w, w_1, \ldots, w_N$ and produces outputs $y_0 = G(w_1), y_1 = G(w_2), \ldots, y_N = G(w_{N+1})$.. (Note that these are precisely the outputs required in order to be matched to the specified tasks.) For lack of a better term, call this an N-*deviation*. From period $N+1$ on the worker will return to producing $G(w)$ forever. Given $w_1, \ldots w_N, w_{N+1}$ define

$$S_N(w_1, \ldots, w_N; w_{N+1}) \triangleq \sum_{t=0}^{N} \delta^t \{\lambda G(w_{t+1}) - c\left[G(w_{t+1})/p(w)q(w_t)\right] k(w)\} \quad (3)$$

This is the (discounted) utility worker w will obtain over this span of periods if the worker follows the given N-deviation.

Given w_{N+1} we study N-deviations $\{w_1, \ldots, w_N, w_{N+1}\}$. Among these, we search for those that are optimal, in the sense of maximizing S_N. To facilitate this search we first study how S_N depends on each of the intermediates $w_1, \ldots, w_N$. We then use that information to show that optimal N-deviations are strictly increasing or strictly increasing or constant. Finally we rule out the first two possibilities, and conclude that optimal N-deviations are constant, From this it will follow quickly that no finite deviation from G can be profitable and hence that G is a SAE, as asserted.

---

that, for worker w, it is always the case that producing output $y > G(w)$ is worse in the current period than producing output y. However, if $y < G(1)$ then producing output $y > G(w)$ in the current period will result in a better task next period and so might be part of a profitable deviation.

To see how S_N depends on w_t note first that w_t appears affects two terms in the sum for S_N: It directly enters the output and cost in period $t-1$ and indirectly enters the cost in period t by affecting the task matched to worker w in period t. More specifically, w_t enters only the following terms in the sum payoff S_N:

$$\begin{aligned}&\delta^{t-1}\left\{P\left[G(w_t)\right]-C\left[\Phi\left(G(w_t),w,w_{t-1}\right),w\right]\right\}\\&\quad+\delta^t\left\{-C\left[\Phi\left(G(w_{t+1}),w,w_t\right),w\right]\right\}.\end{aligned}\tag{4}$$

Note that the only variables that appear in this expression are $w_t; w_{t-1}, w_{t+1}$. Define $Q_w(w_t; w_{t-1}, w_{t+1})$ to be the partial derivative of this expression 4 with respect to w_t. We can calculate it as

$$\begin{aligned}Q_w(w_t;w_{t-1},w_{t+1}) &= \delta^{t-1}\left\{\lambda-\left[\frac{k(w)}{p(w)q(w_{t-1})}\right]c'\left(\frac{G(w_t)}{p(w)q(w_{t-1})}\right)\right\}G'(w_t)\\&\quad-\delta^t\left[\frac{k(w)q'(w_t)G(w_{t+1})}{-p(w)\left[q(w_t)\right]^2}\right]c'\left(\frac{G(w_{t+1})}{p(w)q(w_t)}\right)\end{aligned}$$

We need to analyze the sign of Q_w; to do this it is convenient to do some preliminary work. Define

$$\begin{aligned}D_w(w_t,w_{t-1}) &= \lambda-\left[\frac{k(w)}{p(w)q(w_{t-1})}\right]c'\left(\frac{G(w_t)}{p(w)q(w_{t-1})}\right)\\N_w(w_t,w_{t+1}) &= \left[\frac{k(w)q'(w_t)G(w_{t+1})}{-p(w)\left[q(w_t)\right]^2}\right]c'\left(\frac{G(w_{t+1})}{p(w)q(w_t)}\right)\end{aligned}$$

so that $Q_w = \delta^{t-1}D_wG'(w_t) - \delta^t N_w$. The facts we need about D_w, N_w are summarized in the following lemma.

Lemma 4 *For any w, w_{t-1}, w_t, and w_{t+1}, we have*

(i) $\frac{\delta N_w(w;w)}{D_w(w;w)} = G'(w)$.

(ii) $D_w(w_t; w_{t-1})$ *is strictly increasing in* w*, and strictly increasing in* w_{t-1}*;*

(iii) $N_w(w_t; w_{t+1})$ *is strictly decreasing in* w_{t+1}*.*

(iv) $N_w(w_t; w_{t+1}) < 0$ *for any* w, w_t, w_{t+1}*, and* $D_w(w; w) < 0$*.*

(v) Suppose that $w_{t-1} = w_t = w_{t+1} = \hat{w}$*. Then* $\frac{N_w(\hat{w};\hat{w})}{D_w(\hat{w};\hat{w})}$ *is strictly increasing in* w *in the domain of* w *such that* $D_w(\hat{w}; \hat{w}) < 0$*.*

Proof (i) follows directly from the ODE 1 and the definitions of D_w, N_w.

To obtain (ii), note that, since $q(w_{t-1})$ is increasing in w_{t-1} and since $c'(e)$ is strictly increasing in e, we have $D_w(w_t; w_{t-1})$ is strictly increasing in w_{t-1}. Since $k(w)$ is decreasing in w, $p(w)$ is increasing in w, and $c'(e)$ is strictly increasing in e, we see that $D_w(w_t; w_{t-1})$ is strictly increasing in w.

To obtain (iii), note that since $G(w_{t+1})$ is increasing in w_{t+1} and $c'(e)$ is strictly increasing in e, so $N_w(w_t; w_{t+1})$ is strictly decreasing in w_{t+1}.

To obtain (iv), note that $G(w_{t+1}) \geq G(B) > 0$ for $B > 0$. Combined with our assumptions, this implies that $N_w(w_t; w_{t+1}) < 0$. Since $\frac{\delta N_w(w;w)}{D_w(w;w)} = G'(w)$ and $G'(w) > 0$, we have $D_w(w; w) < 0$.

To obtain (v), it is convenient to study $\frac{D_w(\hat{w};\hat{w})}{N_w(\hat{w};\hat{w})}$, instead of studying $\frac{N_w(\hat{w};\hat{w})}{D_w(\hat{w};\hat{w})}$ directly. We have

$$\begin{aligned}
\frac{D_w(\hat{w};\hat{w})}{N_w(\hat{w};\hat{w})} &= \frac{\lambda - \frac{k(w)}{p(w)q(\hat{w})} c'\left(\frac{G(\hat{w})}{p(w)q(\hat{w})}\right)}{-\frac{k(w)q'(\hat{w})G(\hat{w})}{p(w)[q(\hat{w})]^2} c'\left(\frac{G(\hat{w})}{p(w)q(\hat{w})}\right)} \\
&= -\frac{\lambda\left[q(\hat{w})\right]^2}{q'(\hat{w})G(\hat{w})} \frac{p(w)}{k(w)c'\left(\frac{G(\hat{w})}{p(w)q(\hat{w})}\right)} + \frac{q(\hat{w})}{q'(\hat{w})G(\hat{w})}
\end{aligned}$$

Since $p(w)$ is increasing in w, $k(w)$ is decreasing in w, and $c'(e)$ is strictly increasing in e, we have $\frac{p(w)}{k(w)c'\left(\frac{G(\hat{w})}{p(w)q(\hat{w})}\right)}$ is strictly increasing in w. Therefore, $\frac{D_w(\hat{w};\hat{w})}{N_w(\hat{w};\hat{w})}$ is strictly decreasing in w. Note that $\frac{D_w(\hat{w};\hat{w})}{N_w(\hat{w};\hat{w})}$ can be positive or negative, and that $N_w(\hat{w};\hat{w})$ is always negative. Hence, $\frac{N_w(\hat{w};\hat{w})}{D_w(\hat{w};\hat{w})}$ is strictly increasing in w in the domain of w such that $D_w(\hat{w};\hat{w})$ does not change its sign. This completes the proof. ■

The next lemma isolates the relevant properties of $Q_w(w_t; w_{t-1}, w_{t+1})$.

Lemma 5 *The sign of the derivative $Q_w(w_t; w_{t-1}, w_{t+1})$ satisfies:*

(i) When $w < w_{t+1}$, we have

$$Q_w(w_t; w_{t-1}, w_{t+1}) \begin{cases} > 0, & \text{if } w_t \leq \min\{w, w_{t-1}\} \\ < 0, & \text{if } w_t \geq \max\{w_{t+1}, w_{t-1}\} \end{cases}.$$

(ii) When $w > w_{t+1}$, we have

$$Q_w(w_t; w_{t-1}, w_{t+1}) \begin{cases} > 0, & \text{if } w_t \leq \min\{w_{t+1}, w_{t-1}\} \\ < 0, & \text{if } w_t \geq \max\{w, w_{t-1}\} \end{cases}.$$

(iii) When $w = w_{t+1}$, we have

$$Q_w(w_t; w_{t-1}, w_{t+1}) \begin{cases} > 0, & \text{if } w_t \leq \min\{w_{t+1}, w_{t-1}\} \text{ and } w_t < w_{t+1} \\ < 0, & \text{if } w_t \geq \max\{w, w_{t-1}\} \text{ and } w_t > w \end{cases},$$

Proof An immediate result of (i) of Lemma 4 is that

$$\begin{aligned} Q_w(w_t; w_{t-1}, w_{t+1}) &= \delta^{t-1} \cdot D_w(w_t; w_{t-1}) \cdot \left[\delta \cdot \frac{N_{w_t}(w_t; w_t)}{D_{w_t}(w_t; w_t)}\right] - \delta^t \cdot N_w(w_t; w_{t+1}) \\ &= \delta^t \cdot \left[D_w(w_t; w_{t-1}) \cdot \frac{N_{w_t}(w_t; w_t)}{D_{w_t}(w_t; w_t)} - N_w(w_t; w_{t+1})\right]. \end{aligned}$$

From this equality and the various parts of Lemma 4, we establish Lemma 5 case by case.

(i) Assume $w < w_{t+1}$. Suppose that $w_t \leq \min\{w, w_{t-1}\}$. If $D_w(w_t; w_{t-1}) \geq 0$, since $G(w)$ is increasing and $N_w(w_t; w_{t+1})$ is always negative, we will have $Q_w(w_t; w_{t-1}, w_{t+1}) > 0$. If $D_w(w_t; w_{t-1}) < 0$, we will have $D_{w_t}(w_t; w_t) \leq D_w(w_t; w_t) \leq D_w(w_t; w_{t-1}) < 0$, and hence,

$$\begin{aligned}
Q_w(w_t; w_{t-1}, w_{t+1}) &= \delta^t \cdot \left[D_w(w_t; w_{t-1}) \cdot \frac{N_{w_t}(w_t; w_t)}{D_{w_t}(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&\geq \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_{w_t}(w_t; w_t)}{D_{w_t}(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&\geq \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_w(w_t; w_t)}{D_w(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&> \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_w(w_t; w_t)}{D_w(w_t; w_t)} - N_w(w_t; w_t) \right] \\
&= 0
\end{aligned}$$

In summary, we have $Q_w(w_t; w_{t-1}, w_{t+1}) > 0$ when $w_t \leq \min\{w, w_{t-1}\}$.

Now suppose that $w_t \geq \max\{w_{t+1}, w_{t-1}\}$. Then we have $D_w(w_t; w_{t-1}) \leq D_w(w_t; w_t) \leq D_{w_t}(w_t; w_t) < 0$. Hence, we have

$$\begin{aligned}
Q_w(w_t; w_{t-1}, w_{t+1}) &= \delta^t \cdot \left[D_w(w_t; w_{t-1}) \cdot \frac{N_{w_t}(w_t; w_t)}{D_{w_t}(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&\leq \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_{w_t}(w_t; w_t)}{D_{w_t}(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&< \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_w(w_t; w_t)}{D_w(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&\leq \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_w(w_t; w_t)}{D_w(w_t; w_t)} - N_w(w_t; w_t) \right] \\
&= 0
\end{aligned}$$

This completes the proof of (i).

(ii) Assume $w > w_{t+1}$. Suppose that $w_t \leq \min\{w_{t+1}, w_{t-1}\}$. If $D_w(w_t; w_{t-1}) \geq 0$, we will have $Q_w(w_t; w_{t-1}, w_{t+1}) > 0$. If $D_w(w_t; w_{t-1}) < 0$, we will have $D_{w_t}(w_t; w_t) \leq D_w(w_t; w_t) \leq D_w(w_t; w_{t-1}) < 0$, and hence,

$$\begin{aligned}
Q_w(w_t; w_{t-1}, w_{t+1}) &= \delta^t \cdot \left[D_w(w_t; w_{t-1}) \cdot \frac{N_{w_t}(w_t; w_t)}{D_{w_t}(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&\geq \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_{w_t}(w_t; w_t)}{D_{w_t}(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&> \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_w(w_t; w_t)}{D_w(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&\geq \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_w(w_t; w_t)}{D_w(w_t; w_t)} - N_w(w_t; w_t) \right] \\
&= 0
\end{aligned}$$

Now suppose that $w_t \geq \max\{w, w_{t-1}\}$. Then we have $D_w(w_t; w_{t-1}) \leq D_w(w_t; w_t) \leq D_{w_t}(w_t; w_t) < 0$. Hence, we have

$$\begin{aligned}
Q_w(w_t; w_{t-1}, w_{t+1}) &= \delta^t \cdot \left[D_w(w_t; w_{t-1}) \cdot \frac{N_{w_t}(w_t; w_t)}{D_{w_t}(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&\leq \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_{w_t}(w_t; w_t)}{D_{w_t}(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&\leq \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_w(w_t; w_t)}{D_w(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&< \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_w(w_t; w_t)}{D_w(w_t; w_t)} - N_w(w_t; w_t) \right] \\
&= 0.
\end{aligned}$$

This completes the proof of (ii).

(iii) Finally, assume $w = w_{t+1}$. Suppose that $w_t \leq \min\{w_{t+1}, w_{t-1}\}$. If $D_w(w_t; w_{t-1}) \geq 0$, we will have $Q_w(w_t; w_{t-1}, w_{t+1}) > 0$. If $D_w(w_t; w_{t-1}) < 0$, we will have $D_{w_t}(w_t; w_t) \leq D_w(w_t; w_t) \leq D_w(w_t; w_{t-1}) < 0$, and hence,

$$Q_w(w_t; w_{t-1}, w_{t+1}) = \delta^t \cdot \left[D_w(w_t; w_{t-1}) \cdot \frac{N_{w_t}(w_t; w_t)}{D_{w_t}(w_t; w_t)} - N_w(w_t; w_{t+1}) \right]$$

$$
\begin{aligned}
&\geq \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_{w_t}(w_t; w_t)}{D_{w_t}(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&\geq \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_w(w_t; w_t)}{D_w(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&\geq \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_w(w_t; w_t)}{D_w(w_t; w_t)} - N_w(w_t; w_t) \right] \\
&= 0
\end{aligned}
$$

Note that the second and third inequalities are strict if $w_t < w_{t+1} = w$. Hence, we have $Q_w(w_t; w_{t-1}, w_{t+1}) < 0$ if $w_t < w$ and $w_t \leq \min\{w_{t+1}, w_{t-1}\}$.

Now suppose that $w_t \geq \max\{w, w_{t-1}\}$. Then we have $D_w(w_t; w_{t-1}) \leq D_w(w_t; w_t) \leq D_{w_t}(w_t; w_t) < 0$. Hence, we have

$$
\begin{aligned}
Q_w(w_t; w_{t-1}, w_{t+1}) &= \delta^t \cdot \left[D_w(w_t; w_{t-1}) \cdot \frac{N_{w_t}(w_t; w_t)}{D_{w_t}(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&\leq \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_{w_t}(w_t; w_t)}{D_{w_t}(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&\leq \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_w(w_t; w_t)}{D_w(w_t; w_t)} - N_w(w_t; w_{t+1}) \right] \\
&\leq \delta^t \cdot \left[D_w(w_t; w_t) \cdot \frac{N_w(w_t; w_t)}{D_w(w_t; w_t)} - N_w(w_t; w_t) \right] \\
&= 0
\end{aligned}
$$

Note that the second and third inequalities are strict if $w_t > w_{t+1} = w$. Hence, we have $Q_w(w_t; w_{t-1}, w_{t+1}) < 0$ if $w_t > w$ and $w_t \geq \max\{w, w_{t-1}\}$. This completes the proof of (iii). ■

Using these lemmas, we now show that optimal N-deviations are either monotone or constant. .

Lemma 6 *Let $w^* = \{w_t^*\}_{t=1}^N$ be any solution to the optimization problem*

$$
\max_{\tilde{w}} S_N(\tilde{w}; w_{N+1}) \tag{5}
$$

(i) If $w < w_{N+1}$ *then* $w < w_1^* < \ldots < w_N^* < w_{N+1}$.

(ii) If $w > w_{N+1}$ *then* $w > w_1^* > \ldots > w_N^* > w_{N+1}$.

(iii) If $w = w_{N+1}$ *then* $w = w_1^* = \ldots = w_N^* = w_{N+1}$.

Proof The proof is by induction on N. When $N = 1$, we have $w_{N-1} = w_0 = w$, and aim to find the optimal w_1 given w_2.

(i) When $w < w_2$, part (i) of Lemma 5 tells us that $Q_w(w_1; w_0, w_2) > 0$ when $w_1 \leq w$ and that $Q_w(w_1; w_0, w_2) < 0$ when $w_1 \geq w_2$. Hence, the optimal w_1 lies in (w, w_2).

(ii) When $w > w_2$, part (ii) of Lemma 5 tells us that $Q_w(w_1; w_0, w_2) > 0$ when $w_1 \leq w_2$ and that $Q_w(w_1; w_0, w_2) < 0$ when $w_1 \geq w$. Hence, the optimal w_1 lies in (w_2, w).

(iii) When $w = w_2$, part (iii) of Lemma 5 tells us that $Q_w(w_1; w_0, w_2) > 0$ when $w_1 < w$ and that $Q_w(w_1; w_0, w_2) < 0$ when $w_1 > w_2$. Hence, the optimal w_1 must be w.

This proves the lemma when $N = 1$.

Now suppose Lemma 6 holds at $N-1$; i.e., namely any sequence $\{w_t^*\}_{t=1}^{N-1}$ that maximizes $S_{N-1}(w_1, \ldots, w_{N-1}; w_N)$ is strictly monotone or constant. We want to prove that any sequence $\{w_t^*\}_{t=1}^{N}$ that maximizes $S_N(w_1, w_2, \ldots, w_N; w_{N+1})$ is also strictly monotone or constant. We consider the three cases in turn.

(i) Consider the case with $w < w_{N+1}$.

- Assume that the optimal $w_N^* \leq w$. Then the optimal sequence $\{w_t^*\}_{t=1}^{N-1}$ that maximizes $S_{N-1}(w_1, \ldots, w_{N-1}; w_N)$ must satisfy $w \geq w_1^* \geq \cdots \geq w_{N-1}^* \geq w_N^*$. So we have $w_N^* \leq w_{N-1}^* \leq w < w_{N+1}$. Part (i) of Lemma 5 tells us that $Q_w(w_N^*; w_{N-1}^*, w_{N+1}) > 0$. This is contradictory to the fact that w_N^* is optimal.

- Assume that the optimal $w_N^* \geq w_{N+1}$. Then the optimal sequence $\{w_t^*\}_{t=1}^{N-1}$ that maximizes $S_{N-1}(w_1, \ldots, w_{N-1}; w_N)$ must satisfy $w < w_1^* < \cdots < w_{N-1}^* < w_N^*$. So we have $w_N^* \geq w_{N+1}$ and $w_N^* > w_{N-1}^*$. Part (i) of Lemma 5 tells us that $Q_w(w_N^*; w_{N-1}^*, w_{N+1}) < 0$. This is contradictory to the fact that w_N^* is optimal.

In sum: when $w < w_{N+1}$, the optimal w_N^* must lie in (w, w_{N+1}). Therefore, the sequence $\{w_t^*\}_{t=1}^{N}$ that maximizes $S_N(w_1, w_2, \ldots, w_N; w_{N+1})$ must satisfy $w < w_1^* < \cdots < w_{N-1}^* < w_N^* < w_{N+1}$.

(ii) Consider the case with $w > w_{N+1}$.

- Assume that the optimal $w_N^* \geq w$. Then the optimal sequence $\{w_t^*\}_{t=1}^{N-1}$ that maximizes $S_{N-1}(w_1, \ldots, w_{N-1}; w_N)$ must satisfy $w \leq w_1^* \leq \cdots \leq w_{N-1}^* \leq w_N^*$. So we have $w_N^* \geq w_{N-1}^* \geq w > w_{N+1}$. Part (ii) of Lemma 5 tells us that $Q_w(w_N^*; w_{N-1}^*, w_{N+1}) < 0$. This is contradictory to the fact that w_N^* is optimal.

- Assume that the optimal $w_N^* \leq w_{N+1}$. Then the optimal sequence $\{w_t^*\}_{t=1}^{N-1}$ that maximizes $S_{N-1}(w_1, \ldots, w_{N-1}; w_N)$ must satisfy $w > w_1^* > \cdots > w_{N-1}^* > w_N^*$. So we have $w_N^* \leq w_{N+1}$ and $w_N^* < w_{N-1}^*$. Part (ii) of Lemma 5 tells us that $Q_w(w_N^*; w_{N-1}^*, w_{N+1}) > 0$. This is contradictory to the fact that w_N^* is optimal.

In sum: when $w > w_{N+1}$, the optimal w_N^* must lie in (w_{N+1}, w). Therefore, the sequence $\{w_t^*\}_{t=1}^N$ that maximizes $S_N(w_1, w_2, \ldots, w_N; w_{N+1})$ must satisfy $w > w_1^* > \cdots > w_{N-1}^* > w_N^* > w_{N+1}$.

(iii) Finally, consider the case with $w = w_{N+1}$.

- Assume that the optimal $w_N^* > w$. Then the optimal sequence $\{w_t^*\}_{t=1}^{N-1}$ that maximizes $S_{N-1}(w_1, \ldots, w_{N-1}; w_N)$ must satisfy $w < w_1^* < \cdots < w_{N-1}^* < w_N^*$. So we have $w_N^* > w_{N-1}^* > w = w_{N+1}$. Part (iii) of Lemma 5 tells us that $Q_w(w_N^*; w_{N-1}^*, w_{N+1}) < 0$. This is contradictory to the fact that w_N^* is optimal.
- Assume that the optimal $w_N^* < w$. Then the optimal sequence $\{w_t^*\}_{t=1}^{N-1}$ that maximizes $S_{N-1}(w_1, \ldots, w_{N-1}; w_N)$ must satisfy $w > w_1^* > \cdots > w_{N-1}^* > w_N^*$. So we have $w_N^* < w_{N-1}^* < w = w_{N+1}$. Part (iii) of Lemma 5 tells us that $Q_w(w_N^*; w_{N-1}^*, w_{N+1}) > 0$. This is contradictory to the fact that w_N^* is optimal.

In summary, when $w = w_{N+1}$, the optimal w_N^* must be w. Therefore, the sequence $\{w_t^*\}_{t=1}^N$ that maximizes $S_N(w_1, w_2, \ldots, w_N; w_{N+1})$ must satisfy $w = w_1^* = \cdots = w_{N-1}^* = w_N^* = w_{N+1}$.

This concludes the proof of the lemma. ■

On the basis of these lemmas we can now complete the proof of Theorem 1. If G is not a SAE then some worker w has a profitable deviation. Lemma 2 guarantees that worker w has a profitable finite deviation in which worker w produces the sequence of outputs $G(w_1), G(w_2), \ldots, G(w_N)$ and then produces $G(w)$ forever after. However Lemma 6 tells us that among all

N-deviations that end with $w_{N+1} = w$, the optimal one is constant. That is, there is no profitable deviation, and G is a stationary assortative equilibrium, as asserted. This completes the proof of Theorem 1.

Appendix B: Proof of Theorem 2

To show that $\Pi_{\text{random}} < \Pi_{\text{assort}}$ it is sufficient to show that when workers choose the myopically optimal output the optimal matching is assortative. To show this, it is sufficient (and in fact necessary) to verify that when workers choose the myopically optimal output the imputed output function is supermodular. As we show below, this verification is a straightforward computation.

Write $Y^*(w,t)$ for the output produced when worker w is matched to task t and chooses the myopically optimal effort $e^*(w,t)$. To show that Y^* is supermodular we must show that the mixed partial is positive; to do this we first need to compute partials of e^*.

The myopically optimal effort $e^*(w,t)$ is defined by the first order condition

$$\lambda p(w)q(t) - k(w)c'\left[e^*(w,t)\right] = 0. \tag{6}$$

Implicit differentiation shows that

$$\begin{aligned} \frac{\partial e^*(w,t)}{\partial w} &= \frac{\lambda p'(w)q(t) - k'(w)c'\left[e^*(w,t)\right]}{k(w)c''\left[e^*(w,t)\right]} \\ \frac{\partial e^*(w,t)}{\partial t} &= \frac{\lambda p(w)q'(t)}{k(w)c''\left[e^*(w,t)\right]} \end{aligned}$$

$$\frac{\partial^2 e^*(w,t)}{\partial w \partial t} = \frac{\lambda q'(t)\left[p'(w)k(w) - p(w)k'(w)\right]\left[1 - \frac{c'[e^*(w,t)]c'''[e^*(w,t)]}{\{c''[e^*(w,t)]\}^2}\right]}{[k(w)]^2 \, c''\left[e^*(w,t)\right]}$$

By definition, $Y^*(w,t) = Y\left[e^*(w,t), w, t\right]$, so we differentiate to obtain:

$$\begin{aligned}\frac{\partial^2 Y^*(w,t)}{\partial w \partial t} &= p(w)q(t)\frac{\partial^2 e^*(w,t)}{\partial w \partial t} + p'(w)q'(t)e^*(w,t) \\ &+ p(w)q'(t)\frac{\partial e^*(w,t)}{\partial w} + p'(w)q(t)\frac{\partial e^*(w,t)}{\partial t}.\end{aligned}$$

Our computations of the derivatives of e^* and our assumptions (especially log-concavity of c') guarantee that each of the terms on the right-hand side is positive so we conclude that Y^* is supermodular. As we have noted this guarantees that $\Pi_{\text{random}} < \Pi_{\text{assort}}$.

To see that $\Pi_{\text{assort}} < \Pi_{\text{SAE}}$ consider the ODE (1). Note that the denominator has the form $q(w)F(w)$ and that $F(w) = 0$ is precisely the first-order condition for myopically optimal choice of effort for worker w when matched with task w. At the SAE, $G'(w)$ and the numerator of the ODE are strictly positive so $F(w)$ is also strictly positive; it follows that at the SAE, worker w is exerting *more* than the myopically optimal effort and hence producing output *greater* than $Y^*(w,w)$. Because profit is a fixed fraction of output, the firm obtains greater output from each worker at the SAE than when matching is assortative and workers choose myopically optimal effort. In particular, $\Pi_{\text{assort}} < \Pi_{\text{SAE}}$.

Finally, to see that $\Pi_{\text{SAE}} < \Pi_{\text{FI}}$ note that in the full information setting the firm induces the effort level that leaves the worker with 0 net utility. In the SAE, each worker obtains strictly positive utility. (Any worker who obtained 0 utility could simply produce slightly less output in the current period and exert 0 effort in the future, thereby obtaining strictly positive utility.). Hence, in the full information setting, workers exert greater effort than in SAE and hence produce greater output and greater profit for the firm, so $\Pi_{\text{SAE}} < \Pi_{\text{FI}}$. This completes the proof of Theorem 2.

Appendix C: Counterexample

Finally we keep the promise made in Section 6 to exhibit an example in which, when workers choose myopically optimal effort, random matching may yield greater profit for the firm than assortative matching. In fact we show that if $p(w) = w^\beta$ with $\beta > 0$, $q(t) = t$, $k(w) = \frac{1}{w}$, and $c(e) = 1 - (1-e)^\alpha$ with $\alpha \in (0,1)$ provided that $B \geq \left(\frac{\alpha}{\lambda}\right)^{\frac{1}{3}}$ and that β is small enough, then when worker w is matched with task t and chooses myopically optimal effort (as in the proof of Theorem 2) the imputed output $Y^*(w,t)$ is *submodular*. It follows that, for fixed payment scheme $P(y) = \lambda y$ the *anti-assortative* matching is optimal (yields greatest output and hence greatest firm profit) and that in fact the assortative matching is worst (yields least ouput and hence least firm profit); in particular, assortative matching is worse than random matching so $\Pi_{\text{random}} > \Pi_{\text{assort}}$.

This is a straightforward computation following the same procedure as in the proof of Theorem 2. If worker w is matched to task t and exerts effort e its utility will be

$$\lambda w^\beta te - (1/w)\left[1 - (1-e)^\alpha\right].$$

Because $B \geq \left(\frac{\alpha}{\lambda}\right)^{\frac{1}{3}}$, the myopically optimal effort will be

$$e^*(w,t) = 1 - \left(\frac{\alpha}{\lambda w^{1+\beta} t}\right)^{\frac{1}{1-\alpha}}.$$

Differentiating shows that

$$\frac{\partial^2 Y^*(w,t)}{\partial w \partial t} = \frac{\alpha\beta+1}{w^{\frac{1-\alpha+\alpha\beta}{1-\alpha}}}\left[\frac{\beta w^{\frac{\beta}{1-\alpha}}}{\alpha\beta+1} - \left(\frac{\alpha}{\lambda}\right)^{\frac{1}{1-\alpha}} \frac{\alpha}{(1-\alpha)^2}\left(\frac{w}{t}\right)^{\frac{1}{1-\alpha}}\right]$$

Since $w \leq 1$ and $\left(\frac{w}{t}\right)^{\frac{1}{1-\alpha}} \geq B^{\frac{1}{1-\alpha}}$, we have

$$\frac{\partial^2 Y^*(w,t)}{\partial w \partial t} \leq \frac{\alpha\beta+1}{w^{\frac{1-\alpha+\alpha\beta}{1-\alpha}}}\left[\frac{\beta}{\alpha\beta+1} - \left(\frac{\alpha}{\lambda}\right)^{\frac{1}{1-\alpha}} \frac{\alpha}{(1-\alpha)^2} B^{\frac{1}{1-\alpha}}\right]$$

Since $\left(\frac{\alpha}{\lambda}\right)^{\frac{1}{1-\alpha}} \frac{\alpha}{(1-\alpha)^2} B^{\frac{1}{1-\alpha}}$ is a positive constant, and since $\lim_{\beta\to 0} \frac{\beta}{\alpha\beta+1} = 0$, if $\beta > 0$ is small enough then

$$\frac{\beta}{\alpha\beta+1} < \left(\frac{\alpha}{\lambda}\right)^{\frac{1}{1-\alpha}} \frac{\alpha}{(1-\alpha)^2} B^{\frac{1}{1-\alpha}}.$$

Hence for any such β we see that $\frac{\partial^2 Y^*(w,t)}{\partial w \partial t} < 0$; i.e. Y^* is submodular and assortative matching is worst possible.

References

[1] Becker, G. (1974), "A Theory of Marriage," in *Economics of the Family: Marriage, Children, and Human Capital*, T. Schultz, ed., UMI.

[2] Carroll, G. (2014), "Robustness and Linear Contracts," Stanford University Working Paper.

[3] Cole, H., G. Mailath and A. Postlewaite (2001), "Efficient Noncontractible Investments in Large Economies," *Journal of Economic Theory* 101, 333-373.

[4] Gale, D. and L. Shapley (1962), "College Admissions and the Stability of Marriage," *American Mathematical Monthly* 69, 9-14.

[5] Gretsky, N., J. Ostroy and W. Zame (1999), "Perfect Competition in the Continuous Assignment Model," *Journal of Economic Theory* 88, 60118.

[6] Grossman, G., Helpman, E. and Kircher, P. (2014), "Matching, Sorting, and the Distributional Effects of International Trade," Princeton University Working Paper.

[7] Hopkins, E. (2012), "Job Market Signaling of Relative Position, or Becker Married to Spence," *Journal of the European Economic Association* 10, 290-322.

[8] Mailath, G. (1987), "Incentive Compatibility in Signaling Games with a Continuum of Types," *Econometrica* 55, 1349-1365.

[9] Makowski, L. (2004), "Pre-contractual Investment without the Fear of Holdups: The Perfect Competition Connection," UC Davis Working paper.

[10] Shapley, L. and M. Shubik, "The Assignment Game, I: The Core," *International Journal of Game Theory* 1, 1432-1270

[11] Shimer, R. and L. Smith, "Assortative Matching and Search," *Econometrica* 68, 343-369.