Matching with Evolving Human Capital*

Axel Anderson†
Georgetown University
Economics Department
Washington, DC 20057

Lones Smith‡
University of Michigan
Economics Department
Ann Arbor, MI 48109-1220

(January 2008)
VERY Preliminary: Please do NOT circulate

*We thank seminar participants at the 3rd Annual Cowles-Caress GE Conference, U. Penn Search and Matching Workshop, and Georgetown University.
†Email: aza@georgetown.edu
‡Email: lones@umich.edu; web: www.umich.edu/~lones
1 Introduction

Framework: We introduce a market for partners: matches determine today’s output and influence tomorrow’s human capital.

Questions:

- How does the planner/market rank distributions over human capital?
- What are the returns to human capital acquisition for individuals in the market? Are individuals risk taking when it comes to human capital acquisition?
- How does the introduction of human capital dynamics alter the allocation of partners to matches and the output shares (wages) within matches?
- When are static and dynamic efficiency concerns at odds? Which dominates for high rates of patience?


Outline of Results:

- Existence and Welfare Theorems established in Theorem 0.
- Our answer to the question of how the social planner ranks distributions over human capital is contained in Theorems 1 and Corollary 1. This also characterizes returns to human capital at the individual level.
- Theorem 2 shows that Mentor-Protege transitions imply that PAM can never obtain (for any weights on the present vs. future!).
- Theorem 3 shows that Bad Apple (min) transitions are necessary and sufficient to robustly generate PAM.
- Sufficient conditions for PAM in the deterministic case are contained in Corollary 2.
- Theorem 4 gives necessary and sufficient conditions for PAM to obtain with enough weight on the future in the deterministic transitions case. A corollary follows this result that gives easily interpretable sufficient conditions for PAM and not PAM given enough weight on the future.
- Throughout we provide examples of standard output functions and intuitive transition functions that illustrate the theory.
2 The Model

2.1 Model Basics

- Pairwise matching model with no defined sides.
- Discrete time, infinite horizon model with a continuum of heterogeneous risk neutral agents.
- At the micro level an agent’s type is \( x \in [0,1] \), which evolves stochastically dependent on partner type \( y \). The match transition function, \( \tau(s|x,y) \), is the chance that \( x \) updates to at most \( s \) plus the probability that \( y \) updates to at most \( s \). We assume \( \tau \) is monotonically decreasing in \( x \) and \( y \).
- At the macro level, the distribution over types \( G : [0,1] \rightarrow [0,1] \) is the state variable, which evolves deterministically based on the distribution over matches in the economy, \( F : [0,1]^2 \rightarrow [0,1] \).
- Within each period agent’s choose partners and produce perfectly divisible output, \( Q : [0,1]^2 \rightarrow R_+ \), which we assume monotonic in \( x \) and \( y \).
- Information: There is no private information. There can be symmetric incomplete information as in the Bayesian Reputation Model.
- Timing: Within each period, matches are formed, output is realized, human capital transitions.
- We say that \( x \) and \( y \) are matched if \((x, y)\) lies in the matching set, namely the joint support of \( F(x, y) \). Positive assortative matching (PAM) obtains iff like types match, \( F(x, y) = G(x) \land G(y) \), i.e. matching along the 45 degree line.

2.2 The Planner’s Problem

- State Variable: \( G : [0,1] \rightarrow [0,1] \) a type CDF.
- Choice Variable: \( F : [0,1]^2 \rightarrow [0,1] \) a conditional match CDF.
- Feasibility Constraint: \( \Phi(G) \equiv \{ F : 0 \leq F(1,y) \leq G(y) \ \forall y \} \)
- Flow Payoff: \( \int \int Q(x,y)F(dx,dy) = \int QdF \).
- Discounting: \( \delta = \) time preference, \( \sigma = \) survival rate, \( \beta = \delta \sigma \).
- Inflow: At the start of each period distribution \((1-\sigma)\tilde{G}\) enters the economy.
- Transitions: \((1-\sigma)\tilde{G} + \sigma B(F) = \) updated distribution, with:
  \[
  B(F)(s) = \int \int \tau(s|x,y)F(dx,dy)
  \]
• Bellman Equation:

\[ V(G) = \max_{F \in \Phi(G)} (1 - \delta) \int QdF + \delta V((1 - \sigma)\bar{G} + \sigma B(F)) \]

• Let \( v(x) \) be the shadow value to the planner of relaxing the feasibility constraint at type \( x \), and define \( \psi(x|y) = E[v(x')|x, y] \) and \( \Psi(x, y) = \psi(x|y) + \psi(y|x) \).

• First Order Conditions (complementary slackness):

\[ v(x) + v(y) \geq (1 - \beta)Q(x, y) + \beta \Psi(x, y) \]

\( (x, y) \in \text{supp}(F) \Rightarrow v(x) + v(y) \geq (1 - \beta)Q(x, y) + \beta \Psi(x, y) \)

• Steady State Pareto Optimum: A 4-tuple \((V, v, F, G)\), such that:

1. \((V, v)\) solve the Planner’s problem,
2. \(F\) solves the Bellman Equation given \(G\), and
3. Steady state obtains: \(G = (1 - \sigma)\bar{G} + B(F)\).

2.3 Market Equilibrium

• Wages: \(w(x|y)\) = the wage that agent \(x\) earns if matched with \(y\).

• Values: \(v(x)\) = the maximum discounted sum of wages that \(x\) can earn. We have used the same notation as for the Planner’s shadow value functions, as we shall show these coincide.

• A Steady Competitive Equilibrium is a 4-tuple \((v, w, F, G)\), where:

1. \(F \in \Phi(G)\),
2. \(G = (1 - \sigma)\bar{G} + \sigma B(F)\)
3. \(w(x|y) + w(y|x) = Q(x, y)\)
4. \(v(x) = \max_y[(1 - \delta \sigma)w(x|y) + \delta \sigma \psi(x|y)]\)
5. \((\hat{x}, \hat{y}) \in \text{supp}(F) \Rightarrow \hat{y} \in \arg\max_y[(1 - \delta \sigma)w(\hat{x}|y) + \delta \sigma \psi(x|y)]\)
   \(\hat{x} \in \arg\max_x[(1 - \delta \sigma)w(\hat{y}|x) + \delta \sigma \psi(y|x)]\)
2.4 Existence and Welfare Theorems

Theorem 0 A Steady State Pareto Optimum Exists and the welfare theorems obtain:
(a) If \((v, w, F, G)\) is a Steady State Competitive Equilibrium then \((V = \int vdG, v, F, G)\)
is a Steady State Pareto Optimum.
(b) If \((V, v, F, G)\) is a Steady State Pareto Optimum then \((v, w, F, G)\) is a Steady State
Competitive Equilibrium with:

\[
\begin{align*}
\text{Static Wage} & \quad \text{Dynamic Rent } (y \text{ to } x) \\
W(x|y) &= Q(x, y) - v(y) + \frac{\beta}{1-\beta} [\psi(y|x) - v(y)].
\end{align*}
\]

From Anderson and Smith (2007a), adapted to the current generalization (or perhaps just cited). Relate to Gretsky, Ostroy, and Zame (1992) and Cole, Mailath, and Postlewaite (2001).

Note that we not only prove the welfare theorems, but we provide a mapping between any SSCE and a SSPO, and vice versa. We prove that the value functions for agents in the competitive equilibrium are the equivalent to the shadow values on the feasibility constraint for the Planner. Throughout the paper we critically rely on these mappings, working with the Planner’s Problem and then deriving the implications for matching, values, and wages in the competitive equilibrium, and vice versa.

3 Value Characterization

In this section we characterize the Planner’s value \(V\) and the individual agent’s value \(v\). As the Planner’s Problem can be restated as a contraction mapping, our approach is to use the Planner’s Problem to characterize \(V\) and apply the welfare theorems to characterize the resulting \(v\). These characterizations results are a necessary first step toward determining what matching and wages obtain in equilibrium, since the value of a current match is the sum of the output produced and the continuation value of the updated types resulting from a match.

The Planner chooses a matching to maximize a \((1 - \delta, \delta)\) weighted average of flow payoffs and continuation values. Lorentz (1953) showed that Q PAM, implies that PAM is statically efficient, i.e. solves \(\max_{F \in \Phi(G)} \int QdF\), while Becker (1973) showed that PAM is the market outcome in the static problem with transferable utility and Q SPM. To this static problem we have added dynamics, and seek to understand dynamic efficiency, i.e. what maximizes continuation values:

\[
\max_{F \in \Phi(G)} V((1 - \sigma)\tilde{G} + \sigma B(F)).
\]

Unlike the static problem, the answer depends on an endogenous value function, \(V\), which ranks distributions, \(G\). The transition function, \(\tau\), maps from matching choices, \(F\), to continuation distributions, \(B(F)\). Thus, \(V\) and \(\tau\) together determine an order on
First Order Stochastic Dominance: \( \hat{G} \succ_{\text{FOSD}} G \iff G(s) \geq \hat{G}(s) \ \forall s. \)

Mean Preserving Spread Order: \( \hat{G} \succ_{\text{MPS}} G \iff \int s\hat{G}(s)ds = \int sG(s)ds \) and \( \int_0^z G(s)ds \geq \int_0^z \hat{G}(s)ds \ \forall z. \)

Second Order Stochastic Dominance: \( \hat{G} \succ_{\text{SOSD}} G \iff \int_0^z G(s)ds \geq \int_0^z \hat{G}(s)ds \ \forall z. \)

Increasing Convex Order: \( \hat{G} \succ_{\text{ICX}} G \iff \int_1^z (1 - \hat{G}(s))ds \geq \int_1^z (1 - G(s))ds \ \forall z. \)

We call a function of two variables bi-convex (bi-concave) if it is convex (concave) in both variables individually.

**Theorem 1 (Value Function Characterization)** We have:

(a) \( \hat{G} \succ_{\text{FOSD}} G \Rightarrow V(\hat{G}) > V(G) \) and \( v \) is increasing, and

(b) If \( Q \) and \( \int_0^z \tau(s|x, y)ds \) are bi-convex for all \( z \), then \( \hat{G} \succ_{\text{MPS}} G \Rightarrow V(\hat{G}) > V(G) \) and \( v \) is convex. Bi-concavity assumptions imply \( v \) concave and \( V \downarrow \) in MPS.\(^1\)

\( V \) is a fixed point of the Bellman operator \( T \) defined by:

\[
TV(G) = \max_{F \in \Phi(G)} \left[ (1 - \delta) \int \int Q(x, y)F(dx, dy) + \delta V((1 - \sigma)\hat{G} + \sigma B(F)) \right]
\]

Notice that \( T \) is a monotonic contraction in the sup-norm.

**Proof of Part (a)** Assume that for any \( \hat{G} \succ_{\text{FOSD}} G \) we have \( V(\hat{G}) \geq V(G) \). We claim this implies that \( TV(\hat{G}) > TV(G) \), since \( T \) is a contraction this establishes that the unique fixed point is also increasing in FOSD changes in \( G \).

Let \( M(x|x') \) be a transition function that gives the probability that \( x' \) transitions to \( s \leq x \). For any \( \hat{G} \succ_{\text{FOSD}} G \) there exists such a transition function that satisfies:

\( \hat{G}(x) = \int M(x|x')G(dx') \) and \( M(x|x') = 0 \ \forall \ x < x'. \)

Assume \( F \in \Phi(G) \) is optimal for \( G \), and then form \( \hat{F} \in \Phi(\hat{G}) \):

\[
\hat{F}(x, y) \equiv \int \int M(x|x')M(y|y')F(dx', dy').
\]

\(^1\)Anderson and Smith (2007a) Theorem 3 establishes the same convexity results under a stronger set of assumptions: specifically, \( x_t \) a martingale, \( Q \) bi-linear and strictly SPM, \( \int_0^z \tau(s|x, y)ds \) bi-convex and \( \int_0^z \tau(s|x, x)ds \) convex.
Then since $Q$ is monotonic in both arguments static output is higher under $\hat{F}$ than $F$:

$$\int Qd\hat{F} = \int \int \int M(dx'|x)M(dy'|y)F(dx',dy') > \int QdF$$

Since $\tau$ is monotonically decreasing in $x$ and in $y$, we have:

$$B(\hat{F})(s) = \int \int \tau(s|x,y)\hat{F}(dx,dy) = \int \int \int \tau(s|x,y)M(dx|x')M(dy|y')F(dx',dy')$$

$$\leq \int \int \tau(s|x',y')F(dx',dy') = B(F)(s)$$

Thus, $B(\hat{F}) \succ_{FOSD} B(F)$, which in turn implies

$$V(\sigma B(\hat{F}) + (1 - \sigma)\bar{G}) \geq V(\sigma B(F) + (1 - \sigma)\bar{G}).$$

Altogether, $TV(\hat{G}) > TV(G)$.

**Proof of Monotonicity of $v$.** For any $x$ in the support of $G$, increase a small mass $\varepsilon$ of the $G$ distribution near $x$ to $x + h$, where $h > 0$ is feasible and arbitrary. The slope of $V(G) = \int v(x)G(dx)$ in $\varepsilon$ is proportional to $v(x + h) - v(x)$, at $\varepsilon = 0$. Since $V(G)$ rises in any FOSD shift of $G$, this must be strictly positive. So $v(x)$ is everywhere increasing.

**Proof of Part (b)** Assume that for any $\hat{G} \succ_{MPS} G$ we have $V(G) \geq V(\hat{G})$. We claim this implies that $TV(\hat{G}) > TV(G)$.

Let $M(x|x')$ be a transition function that gives the probability that $x'$ transitions to $s \leq x$. For any $\hat{G} \succ_{MPS} G$ there exists such a transition function that satisfies:

$$\hat{G}(x) = \int M(x|x')G(dx') \quad \text{and} \quad \int xM(dx|x') = x' \quad \forall \ x'.$$

Assume $F \in \Phi(G)$ is optimal for $G$, and let $\hat{F} \in \Phi(\hat{G})$ satisfy:

$$\hat{F}(x,y) = \int \int M(x|x')M(y|y')F(dx',dy').$$

Then since $Q$ is bi-convex static output is higher under $\hat{F}$ than $F$:

$$\int Qd\hat{F} = \int \int \int Q(x,y)M(dx|x')M(dy|y')F(dx',dy')$$

$$> \int \int Q \left( \int xM(dx|x')dx, \int yM(dy|y')dy \right) F(dx',dy')$$

$$= \int \int Q(x',y')F(dx',dy') = \int QdF.$$
Likewise for continuation distributions we have:

\[
\int_0^z B(\hat{F})(s) ds = \int_0^z \int \int \tau(s|x, y) \hat{F}(dx, dy) ds \\
= \int \int \int \int \int_0^z \tau(s|x, y) M(dx|x') M(dy|y') F(dx', dy') \\
\geq \int \int \int \int_0^z \tau(s|x', y') F(dx' dy') = \int_0^z B(F)(s) ds
\]

where the first inequality follows from \( \int_0^z \tau(s|x, y) ds \) bi-convex. So that, we have \( B(\hat{F}) \succ MPS B(F) \), which in turn implies

\[
V(\sigma B(\hat{F}) + (1 - \sigma) \bar{G}) \geq V(\sigma B(F) + (1 - \sigma) \bar{G}).
\]

Altogether, \( TV(\hat{G}) > TV(G) \).

**Proof of Convexity of \( v \).** For any \( x \) in the support of \( G \), equally spread a small fraction \( \varepsilon \) of the \( G \) distribution near \( x \) to \( x \pm h \), where \( h > 0 \) is feasible and arbitrary. The slope of \( V(G) = \int v(x)G(dx) \) in \( \varepsilon \) is proportional to \([v(x + h) + v(x - h)]/2 - v(x)\), at \( \varepsilon = 0 \). Since \( V(G) \) rises in any MPS of \( G \), this must be strictly positive. So the planner’s shadow value \( v(x) \) is everywhere convex.

Clearly we may reverse the inequality when \( Q \) and \( \int_0^z \tau \) are bi-concave. \( \Box \)

The assumptions required for \( V \) increasing in FOSD changes are quite mild: higher types produce more and one’s type tomorrow is probabilistically increasing in one’s type today.

We can relate on preferences for spread to the small literature on efficient macroeconomic inequality, ex. Atkeson and Lucas (1992) and Welch (1999). Although both are related in that they argue there for inequality in an efficient equilibrium, the mechanisms that make inequality efficient are quite different. In Atkeson and Lucas (1992) agents have private information and inequality must obtain to make revealing this private information incentive compatible. Welch (1999) is mostly an empirical piece that draws a distinction between unequal outcomes and unequal opportunity. He argues that inequality of outcomes can provide incentives for effort, education, etc., as long as opportunities are not too unequal.

**Corollary 1** Given \( Q \) and \( \int_0^z \tau(s|x, y) ds \) bi-convex: \( \hat{G} \succ ICX G \Rightarrow V(\hat{G}) > V(G) \). Conversely, bi-concavity assumptions imply \( V \uparrow \) in SOSD shifts in \( G \).

**Proof:** By Shaked and Shanthikumar (2007) Theorem 4.A.6, \( \hat{G} \succ SOSD G \) implies that there exists a \( G' \) such that: \( G' \succ FOSD G \) and \( G' \succ MPS \hat{G} \). If instead \( \hat{G} \succ ICX G \) then \( \exists G' \) such that: \( G' \succ FOSD G \) and \( \hat{G} \succ MPS G' \). Together these establish the result. \( \Box \)

In the next section we explore the implications of these value characterizations in a world with deterministic transitions.
4 Deterministic Transitions

Consider dynamic efficiency concerns with deterministic transitions. Specifically if \(x\) and \(y\) match then \(x\) transitions to \(f(x, y)\) and \(y\) transitions to \(f(y, x)\).\(^2\) Given this specification for transitions, an individual agent’s expected continuation value function becomes \(E[v(x_{t+1}|x_t, y_t)] = \psi(x_t|y_t) = v(f(x_t, y_t))\).

To satisfy our monotonicity assumption on \(\tau\), \(f\) must be weakly increasing. We assume further that \(x \leq y \Rightarrow f(x, y) \leq f(y, x)\) and that human capital evolution is bounded between the min and max of partner types: \(f(x, y) \in [\min\{x, y\}, \max\{x, y\}]\).

4.1 Two Critical Archtypes

Consider the following two extreme cases:

**Mentor/Protege:** \(x_{t+1} = y_{t+1} = \max\{x_t, y_t\}\): the agent with lower human capital (the protege) takes on the human capital of his match partner (the mentor).

**Bad Apple:** \(x_{t+1} = y_{t+1} = \min\{x_t, y_t\}\): the agent with lower human capital (the bad apple) pulls down the human capital of his match partner.

Clearly these are both extreme assumptions. However, each plays a critical theoretical role. Further we feel they are archetypal representations of certain human capital interactions. Later we shall return to more general assumptions.

PAM is never Efficient with Mentor/Protege Transitions. One can immediately show that \(v\) increasing implies that the expected continuation value function \(v(f(x, y))\) is SBM, thus PAM minimizes continuation values. We have:

\[
\psi(x|y) = E[v(x'|x, y)] = v(\max\{x, y\}) \equiv v(x \lor y) \Rightarrow \\
v(x^H \lor y^H) + v(x^L \lor y^L) \leq v(x^H \lor y^L) + v(x^L \lor y^H) \forall x^L \leq x^H, y^L \leq y^H.
\]

However we can prove an even stronger result:

**Theorem 2** Given Mentor/Protege transitions and \(Q\) increasing and \(C^2\), PAM cannot obtain in equilibrium.

\(^2\)The basic insights in this section obtain under separable noise:

\[
x_{t+1} = (1 - \varepsilon)f(x_t, y_t) + \varepsilon Z(x_t) \\
y_{t+1} = (1 - \varepsilon)f(y_t, x_t) + \varepsilon Z(y_t),
\]

where \(Z(x)\) is a random variable with conditional distribution \(\xi(z|x)\).
Note that we have made no assumptions about the survival rate and discount factor. In fact, our result is even stronger than stated: in equilibrium there can be no interval \((x, x + \varepsilon)\) such that assortative matching obtains along this interval, as the following proof should make clear.

**Proof:** If assortative matching obtains, then values must solve: 

\[
v(x) = (1 - \beta)Q(x, x)/2 + \beta v(\max\{x, x\}),
\]

so that \(v(x) = Q(x, x)/2\), and the sum of the values from an \((x, x)\) match is simply \(Q(x, x)\). Now assume pairs \((x, x)\) and \((x + \varepsilon, x + \varepsilon)\) match for any \(x\) and \(\varepsilon > 0\), the sum of the values from these matches are \(Q(x, x) + Q(x + \varepsilon, x + \varepsilon)\). If we instead deviated for one period only and formed two non assortative matched pairs \((x + \varepsilon, x)\), the net change in the sum of the values would be:

\[
\Delta(\varepsilon) \equiv 2\left[(1 - \beta)Q(x + \varepsilon, x) + \beta Q(x + \varepsilon, x + \varepsilon)\right] - [Q(x, x) + Q(x + \varepsilon, x + \varepsilon)],
\]

which after some algebraic manipulation yields:

\[
\Delta(\varepsilon) = Q(x + \varepsilon, x) - Q(x, x) + (2\beta - 1)\left[Q(x + \varepsilon, x + \varepsilon) - Q(x + \varepsilon, x)\right].
\]

Now substitute the following Taylor series approximations:

\[
Q(x + \varepsilon, x) - Q(x, x) \approx Q_1(x, x)\varepsilon \quad \text{and} \quad Q(x + \varepsilon, x + \varepsilon) - Q(x + \varepsilon, x) \approx Q_2(x, x)\varepsilon + Q_{12}(x, x)\varepsilon^2
\]

to get:

\[
\Delta(\varepsilon) \approx Q_1(x, x)\varepsilon + (2\beta - 1)\left[Q_2(x, x)\varepsilon + Q_{12}(x, x)\varepsilon^2\right].
\]

Given symmetry: \(Q_1(x, x) = Q_2(x, x)\), this becomes:

\[
\Delta(\varepsilon) = 2\beta Q_1(x, x)\varepsilon + (2\beta - 1)Q_{12}(x, x)\varepsilon^2.
\]

Now fix \(\beta > 0\). By continuity, there exists small enough \(\varepsilon^* > 0\) such that \(\Delta(\varepsilon) > 0\) for all \(\varepsilon \leq \varepsilon^*\). Thus, small one period deviations from assortative matching increase the sum of values and assortative matching cannot obtain in equilibrium. \(\square\)

While we cannot have matching along the 45 degree line, matching must be increasing (when continuous), as long as \(Q\) is strictly SPM. To see why, consider, the match value function

\[
\Gamma(x, y; \beta) \equiv (1 - \beta)Q(x, y) + \beta \Psi(x, y)
\]

which is locally strictly SPM away from the 45 degree line, i.e. for all \(\beta < 1\) and \(x \neq y\) there exists \(\varepsilon > 0\) such that:

\[
\Gamma(x, y; \beta) + \Gamma(x + \varepsilon, y + \varepsilon; \beta) > \Gamma(x, y + \varepsilon; \beta) + \Gamma(x + \varepsilon, y; \beta).
\]

which implies that the optimal matching cannot be continuous and decreasing. Together, the fact that we cannot match along the 45 degree line and that the matching
must be increasing away from the 45 degree line implies that the matching must be discontinuous at some point.

In Figure 1 we solve for the matching set and competitive equilibrium wages in a two period model when $\beta = 0.6$ (remove two period assumption and simply solve the infinite horizon model). Notice that in equilibrium there are two classes of agents, those above $x = 1/2$ (mentors) and those below (proteges). The model produces endogenous negative assortative matching across the two classes: mentors only match with proteges and vice versa, but there is positive assortative matching across matches: mentors with higher human capital match with proteges with higher human capital. This might be a fun model to explore within an overlapping generations framework.

In Figure 2 we graph the matching set for several different values of $\beta$ in the same two period version of the model. Notice that in every case there are subsets of agents for which the same endogenous mentor/protege pattern obtains: negative assortative matching of mentors to proteges but positive assortative matching across matches: proteges with higher human capital match with mentors with higher human capital. However, for lower values of $\beta$, having higher human capital does not make one a mentor. Relate the basic result of ranges of types forming endogenous “classes” to Burdett and Coles (1997).

**Bad Apple Transitions and PAM.** In a static world, $Q$ SPM forces PAM. If human capital evolution is independent of partner choice, i.e. $f(x, y) = x$, then PAM again obtains. In this section we ask: what assumptions on deterministic transitions robustly generate PAM when output is monotonic and SPM? The answer is Bad Apple transitions. More specifically we show: min transitions are sufficient for PAM for any $\beta$ as long as output is SPM; that the interactive human capital effect must be of the
form \( \min\{x, y\} \) for PAM to obtain for all \( Q \) increasing and SPM; and with enough patience \( \min \) transitions alone yield PAM: we can drop the assumption that output is SPM.

Intuitively when transitions are of the form “the low type, pulls the high type down,” PAM is dynamically optimal. What is more surprising is that the extreme Bad Apple transitions are is necessary for PAM to maximize continuation values for all monotonic \( Q \).
We will consider several alternative sets of assumptions for \( Q : [0,1]^2 \rightarrow \mathbb{R} \).

- Let \( Q = \{ Q : Q : [0,1]^2 \rightarrow R, \text{increasing} \} \).
- Let \( Q_+ = \{ Q : Q \in Q, \text{strictly increasing} \} \).
- Let \( Q_{SM} = \{ Q : Q \in Q, \text{SPM} \} \).

**Theorem 3 (Min Transitions)** We have:

(a) If \( Q \in Q_{SM} \) and \( f(x,y) = \min\{x,y\} \), then PAM is optimal \( \forall \beta \).

(b) If PAM is optimal \( \forall \beta \) and \( \forall Q \in Q_{SM} \), then \( f(x,y) = \lambda x + (1 - \lambda) \min\{x,y\} \) for some \( \lambda \in [0,1] \).

(c) If \( Q \in Q_+ \) and \( f(x,y) = \min\{x,y\} \), \( \exists \beta^* \) \( s.t. \) PAM obtains \( \forall \beta > \beta^* \).

**Proof Preliminaries:** Proof of (a): Let \( Q \in Q_{SM} \subset Q \) and \( f \) weakly increasing implies \( v \) increasing, which in turn implies \( \Psi(x,y) = 2v(\min\{x,y\}) \) is SPM, since

\[
v(\min\{x^H, y^H\}) + v(\min\{x^L, y^L\}) \geq v(\min\{x^H, y^L\}) + v(\min\{x^L, y^H\}),
\]

for all \( x^L \leq x^H, y^L \leq y^H \) and \( v \) increasing. Given \( Q \) and \( \Psi \) SPM, \( \Gamma(x,y;\beta) \) is SPM for all \( \beta \), which proves the result as \( \Gamma \) SPM in \( (x,y) \) is a sufficient condition for PAM. \( \square \)

Proof of (b): Assume PAM obtains, then \( v \) satisfies:

\[
2v(x) = (1 - \beta)Q(x,x) + 2\beta v(f(x,x)) = (1 - \beta)Q(x,x) + 2\beta v(x) \Rightarrow 2(1 - \beta)v(x) = (1 - \beta)Q(x,x) \Rightarrow v(x) = Q(x,x)/2.
\]

Given \( v(x) = Q(x,x)/2 \), matching any pair \( (x,y) \) yields total value:

\[
\Gamma(x,y;\beta) = (1 - \beta)Q(x,y) + \beta [Q(f(x,y), f(x,y))/2 + Q(f(y,x), f(y,x))/2].
\]

PAM will not be optimal if \( \Gamma(x,y;\beta) > v(x) + v(y) \), or:

\[
(1-\beta)Q(x,y)+\beta [Q(f(x,y), f(x,y))/2 + Q(f(y,x), f(y,x))/2] > Q(x,x)/2+Q(y,y)/2.
\]

Since the LHS is continuous in \( \beta \), \( \Gamma(x,y;1) > v(x) + v(y) \Leftrightarrow \)

\[
Q(f(x,y), f(x,y)) + Q(f(y,x), f(y,x)) > Q(x,x) + Q(y,y)
\]

for some \( (x,y) \) pair is sufficient to rule out PAM. We will show that \( f(x,y) \neq \min\{x,y\} \) implies that there exists \( Q \in Q_{SM} \) such that this inequality obtains.

Given any increasing function \( h : [0,1] \rightarrow \mathbb{R} \) we can construct an increasing and SPM \( Q \) that obeys \( Q(x,x) = h(x) \). Thus, if we can show that there exists increasing \( h \) and \( (x,y) \) such that:

\[
h(f(x,y)) + h(f(y,x)) > h(x) + h(y)
\]
for some \((x, y)\) whenever \(f(x, y) \neq \lambda x + (1 - \lambda) \min \{x, y\}\) for some \(\lambda \in [0, 1]\) we are done.

WLOG assume \(x < y\). If \(f(x, y) \neq \lambda x + (1 - \lambda) \min \{x, y\}\) for some \(\lambda \in [0, 1]\) then \(\min \{x, y\} \leq f(x, y)\) implies \(x < f(x, y)\), and so since \(x \leq y \Rightarrow f(x, y) \leq f(y, x)\) we must have \(x < f(x, y) \leq f(y, x) \leq y\). Now choose the step function: \(h(x') = 0 \forall x' \leq x\) and \(h(x') = 1\) otherwise, and we have:

\[
h(f(x, y)) + h(f(y, x)) = 2 > 1 = h(x) + h(y),
\]
so Inequality (2) is satisfied, and PAM cannot obtain for high enough \(\beta\).

**Proof of (c):** If \(Q \in \mathcal{Q}_+\) and \(f(x, y) = \min \{x, y\}\), then for \(\beta\) high enough we can construct a competitive equilibrium with PAM, by assuming \(v(x) = Q(x, x)/2\), and then showing that agent’s will choose to match assortatively.

Given \(v(x) = Q(x, x)/2\), the sum of any matched pair \((x, y)\) is:

\[
\Gamma(x, y; \beta) = (1 - \beta)Q(x, y) + \beta Q(\min \{x, y\}, \min \{x, y\}),
\]
Agent’s will choose to match assortatively if:

\[
\Gamma(x, x; \beta) + \Gamma(y, y; \beta) \geq \Gamma(x, y; \beta) + \Gamma(y, x; \beta) \quad \forall x, y.
\]
Given \(\Gamma\) continuous in \(\beta\), this will be satisfied for high \(\beta\) if:

\[
\Gamma(x, x; 1) + \Gamma(y, y; 1) \geq \Gamma(x, y; 1) + \Gamma(y, x; 1) \iff Q(x, x) + Q(y, y) > 2Q(\min \{x, y\}),
\]
which follows from \(Q\) strictly increasing.

To gain some insight into the results for Mentor-Protege and Bad Apple transitions, consider the general continuation value function \(\psi(x|y)\). As in the static world, \(\psi(x|y)\) SPM will imply PAM dynamically efficient. Of course, \(\psi(x|y)\) depends on the endogenous value function \(v\). If we assume \(v\) twice differentiable (to build intuition):

\[
\psi_{xy}(x|y) = v'(f(x, y))f_{xy} + v''(f(x, y))f_x(x, y)f_y(x, y). \tag{3}
\]
We know \(v' > 0\) and \(f_x f_y \geq 0\). To conclude that PAM is dynamically optimal based on this information alone would require \(f_x f_y = 0\) and \(f_{xy} > 0\). As is well known, Bad Apple transitions \(f(x, y) = \min \{x, y\}\) are the limit as \(\rho \to -\infty\) of the CES transition function \(f(x, y) = (\alpha x^\rho + (1 - \alpha y^\rho)^{1/\rho}\), which implies:

\[
f_x f_y = \alpha(1 - \alpha) \left(\frac{f(x, y)^2}{xy}\right)^{1-\rho} \Rightarrow \lim_{\rho \to -\infty} f_x f_y = \lim_{\rho \to -\infty} \alpha(1 - \alpha) \left(\frac{\min \{x, y\}^2}{xy}\right)^{1-\rho} = 0
\]

since \(\min \{x, y\}^2 < xy\). Further we have: \(f_{xy} > 0\) for the CES transition function if \(\rho < 1\): altogether, PAM is dynamically optimal given Bad Apple transitions. On the other extreme is Mentor/Protege transitions \(f(x, y) = \max \{x, y\}\), the limit of the CES transition function as \(\rho \to \infty\). Thus for M-P transitions, \(f_x f_y = 0\) and \(f_{xy} = -\infty\) along the 45 degree line, which implies no \((x, x)\) pairs form in equilibrium.
4.2 Away from the Extreme Archetypes

Now we move away from the extreme M-P and Bad Apple transitions to explore more intermediate cases in which \( v'' \) matters. First we go to another extreme and consider a transition function in which \( v' \) does not influence dynamic preferences over matches.

**Linear Transitions and Preference for Spread.** Consider the linear transition function

\[
 f(x, y) = \alpha x + (1 - \alpha) y,
\]

which captures the idea that one’s human capital can be both negatively and positively influenced by partners. Downward movements could result from picking up bad habits or learning incorrect production techniques or other peer effects. This linear transition function is between M-P and Bad Apple transitions in the sense that all three belong to the CES class, with coefficient \( \rho \) ordered \(-\infty < 1 < \infty\) for Bad Apple, linear, and M-P transitions. While the linear transition function is intermediate between M-P and Bad Apple within the class of CES transition functions it is extreme within our theory: \( f_{xy} = 0 \) and thus dynamic efficiency follows from \( v'' \) alone (\( v' \) does not matter).

To sign \( v'' \), evaluate:

\[
\int_0^z \tau(s|x,y)ds = \max\{0, z - \alpha x - (1 - \alpha)y\} + \max\{0, z - (1 - \alpha)x - \alpha y\}.
\]

Since SPM is preserved by addition and the max operator, \( \int_0^z \tau(s|x,y)ds \) is SPM. Convexity is also preserved by the max operator, so that \( \int_0^z \tau(s|x,y)ds \) is convex along the diagonal \( y = x \). Thus, if \( Q(x,x) \) is convex and SPM (ex. \( Q(x,y) = xy \)), we have \( V \) increasing in mean preserving spreads and \( v \) convex by Theorem 1. In turn, \( v'' > 0 \), implies PAM dynamically optimal.

Of course, the sign of \( v'' \) depends on both static output \( Q \) and transitions. With linear transitions, PAM is only optimal when the Planner favors spread. If instead, \( V \) falls in MPS, then NAM would maximize continuation values. Consider \( Q(x,x) \) concave. Assume PAM does obtain, then \( v(x) = Q(x,x)/2 \), and the Planner’s Value \( V(G) = \int Q(x,x)G(dx)/2 \) will be decreasing in MPS, which in turn implies that PAM minimizes continuation values (NAM is dynamically optimal).

Thus, for \( Q(x,x) \) concave, the optimal matching depends on the tradeoff between static and dynamic matching concerns. It turns out, that as long as \( Q \) is twice smooth dynamic concerns dominate for high \( \beta \) and we cannot have PAM. In fact, as in the M-P transition case we cannot have any \( (x,x) \) pairs matched. To see this result, evaluate the cross partial of match values:

\[
\Gamma_{xy}(x,x;\beta) = (1 - \beta)Q_{xy}(x,x) + \beta\Psi_{xy}(x,x) \\
= (1 - \beta)Q_{xy}(x,x) + 2\beta\alpha(1 - \alpha)(Q_{xx}(x,x) + Q_{yy}(x,x) + 2Q_{xy}(x,x)).
\]

Given differentiability, \( Q(x,x) \) strictly concave, iff \( Q_{xx}(x,x) + Q_{yy}(x,x) + 2Q_{xy}(x,x) < 0 \), thus \( \Gamma_{xy}(x,x;\beta) < 0 \) for \( \beta \) high enough. With \( \Gamma \) SBM along the diagonal \( y = x \) we cannot have \( x \) matched to \( x \) at any \( x \).
For example if $Q(x, y) = 2(xy)^{1/4}$, then $v(x) = \sqrt{x}$, which in turn yields:

$$\psi(x|y) = \sqrt{\alpha x + (1 - \alpha)y} \Rightarrow \Psi_{xy}(x, y)/2 = -\frac{\alpha(1 - \alpha)}{2(\alpha x + (1 - \alpha)y)^{3/2}} < 0.$$ 

PAM cannot be globally optimal for high $\beta$, but neither can NAM. To see why, notice that $\psi_{xy}(x, y)$ is bounded as long as either $x$ or $y$ is positive, while $Q_{xy}(x, y) = (xy)^{-3/4}/8$ tends to infinity if either $x$ or $y$ tends to zero. Thus, near $(x, y) = (0, 1)$ match values are unboundedly SPM for any $\beta < 1$ and NAM cannot hold near $(0, 1)$. Symmetrically NAM cannot hold near $(1, 0)$.

Altogether, match values are SBM along the diagonal for $\beta > 1/2$ and SPM near $(0, 1)$ and $(1, 0)$ for any $\beta < 1$. In Figure 3 we graph the matching set, wages $w(x|y(x))$ and values $v$ for the two period Linear model of transitions with $\alpha = 1/2, Q(xy)^{1/4}$, and $\beta = 0.7$ assuming a uniform distribution over types in the first period.

**Figure 3:** Matching set, 2nd period values, and equilibrium wages ($w(x|y(x))$) for the Linear transitions example with two periods, $\alpha = 1/2, Q(x, y) = 2(xy)^{1/4}$, and $\beta = 0.7$.

**Balancing First and Second Order Preferences** Intuitively, complementarity in transitions $f_{xy} > 0$ should favor PAM, which follows from Equation (3) by $v' > 0$ (established in Theorem 1). But as we have seen this is not the full story: $v'' f_x f_y$ matters.\(^3\) While $f$ is exogenous, the sign of $v''$ depends on properties of $Q$ and $\tau$. In the deterministic case:

$$\int_0^z \tau(s|x, y)ds = \max\{0, z - f(x, y)\} + \max\{0, z - f(y, x)\}.$$ 

\(^3\)While $f_x f_y \geq 0$ seems like a reasonable assumption, one can imagine stories for which $f_x > 0$ and $f_y < 0$. For example, perhaps human capital transitions are the result of learning by doing and lower ability types free ride on higher ability types. Then matching with a higher type today may lead to lower human capital in the future.
From this it follows that $f$ bi-concave implies $\int_0^s \tau(s|x, y)ds$ bi-convex, which via Equation (3) yields the following corollary to Theorem 1.

**Corollary 2** Assume SPM, bi-concave, and monotonic deterministic transition function $f$ and bi-convex output $Q$, then PAM obtains for all $\beta$.

**4.3 When the Future Looms Large**

We have considered several examples in which static and dynamic efficiency conflict and shown that dynamic efficiency concerns dominate in these examples for high $\beta$ in the first period of a two period model. We are interested in how this tradeoff between static and dynamic efficiency plays out for high $\beta$ more generally.

Equation (3) establishes whether dynamic matching concerns favor assortative matching or not. Intuitively, if dynamic matching concerns do not favor PAM and we put enough weight on the future, then dynamic concerns should dominate and PAM should not obtain in equilibrium. However, this logic ignores the fact that value functions are endogenous to the discount and survival rates. It could be that as we place more weight on the future, the absolute value of the first and/or second derivatives of $v$ fall and static matching concerns dominate for all $\beta$. Thus, we must be careful when comparing how the relative importance of static and dynamic matching considerations change in $\beta$. To make the dependence of values on $\beta$ explicit we subscript values with the product of the discount and survival rate: $\beta = \sigma \beta$, $v_{\beta}(x)$ and $V_{\beta}(G)$.

Define the continuation output function: $H : [0, 1]^2 \rightarrow \mathbb{R}$ as:

$$H(x, y) = Q(f(x, y), f(x, y)).$$

Let $H$ satisfy *Single Crossing (SC)* iff for all $x^L < x^H$ and $y^L < y^H$:

$$H(x^L, y^H) \geq H(x^L, y^L) \Rightarrow H(x^H, y^H) > H(x^L, y^L).$$
**Theorem 4** PAM obtains for high $\beta$ iff $H$ satisfies (SC).

**Proof Preliminaries:** Recall the match value function:

$$\Gamma(x, y; \beta) \equiv (1 - \beta)Q(x, y) + \beta \Psi(x, y; \beta), \quad \Psi(x, y; \beta) = v_\beta(f(x, y)) + v_\beta(f(y, x)).$$

Milgrom and Shannon (1994) showed that $Y(x) = \arg \max_y \Gamma(x, y; \beta)$ is an increasing correspondence iff $\Gamma$ satisfies a weak version of our (SC). Symmetry of $\Gamma(x, y; \beta)$ implies that $x \in Y(x)$ for all $x$, when $Y(x)$ is increasing. Then, $Y(x) = x$ follows from the fact that our single crossing assumption is strict.

\[ \sim H \text{ Satisfies (SC)} \Rightarrow \sim \text{PAM for High } \beta: \] Assume PAM obtains, then $v(x) = Q(x, x)/2$, which critically does not depend on $\beta$. Thus,

$$\lim_{\beta \to 1} \Gamma(x, y; \beta) = \frac{1}{2}H(x, y) + \frac{1}{2}H(y, x), \quad (4)$$

and if $H$ does not satisfy (SC), $\Gamma(x, y; \beta)$ will not satisfy (SC) for high $\beta$, and PAM will not obtain by Milgrom and Shannon (1994).

\[ H \text{ Satisfies (SC)} \Rightarrow \text{PAM Obtains for High } \beta: \] We shall construct a competitive equilibrium for high $\beta$. Assume $v(x) = Q(x, x)/2$, then each individual $x$ chooses a partner $y$ to solve:

$$\max_y (\Gamma(x, y; \beta) - v(y)).$$

By (4), $\Gamma(x, y; \beta)$ satisfies (SC) for high $\beta$ iff $H$ satisfies (SC) given $v(x) = Q(x, x)/2$. Further, $\Gamma(x, y; \beta)$ satisfies (SC) implies $\Gamma(x, y; \beta) - v(y)$ satisfies (SC) for high $\beta$. Thus, $x \in \arg \max_y (\Gamma(x, y; \beta) - v(y))$ for high $\beta$, and PAM is a competitive equilibrium. □

**Corollary 3** The following sufficient conditions follow from Theorem 4:

(a) If $Q(x, x)$ is convex and $f$ strictly SPM, then PAM obtains for high $\beta$.

(b) If $Q(x, x)$ is concave and $f$ strictly SBM, with at least one strict, then PAM will not obtain for high $\beta$.

**Proof of (a):** Given $f$ increasing, $Q$ strictly increasing, $Q(x, x)$ convex, and $f$ strictly SPM, $H$ is strictly SPM, which is a sufficient condition for $H$ satisfying (SC). For intuition, consider the differentiable case:

$$H_{12}(x, y) = Q''(f(x, y), f(x, y))f_1(x, y)f_2(x, y) + Q'(f(x, y), f(x, y))f_{12}(x, y) > 0$$

**Proof of (b):** Given $f$ increasing, $Q$ strictly increasing, $Q(x, x)$ concave, and $f$ strictly SBM, $H$ is SBM along the diagonal:

$$H(x, x) + H(y, y) < H(x, y) + H(y, x)$$

□
for all \((x, y) \in [0, 1]^2\), i.e. \(H_{12}(x, x) < 0\) for all \(x\) given twice differentiable \(H\). This in turn implies that \(H\) cannot satisfy (SC). For intuition consider the differentiable case, and evaluate \(H_{12}\) along the diagonal:

\[
H_{12}(x, x) = Q''(x, x)f_1(x, x)f_2(x, x) + Q'(x, x)f_{12}(x, x) < 0.
\]

\(\Box\)

Example: CES Transitions  Let transitions and output be given by:

\[
f(x, y) = (\alpha x^\rho + (1 - \alpha)y^\rho)^{1/\rho}\quad \text{and} \quad Q(x, y) = (\frac{1}{2}x^a + \frac{1}{2}y^a)^b.
\]

Then we may calculate:

\[
H_{12}(x, y) = (ab - \rho)a\alpha(1 - \alpha)(xy)^{\rho-1}f(x, y)^{ab-2\rho},
\]

and \(H\) is SPM (SBM) as \(ab > \rho\) \((ab < \rho)\). So that PAM will obtain for high \(\beta\) iff \(ab > \rho\).

5 Stochastic Transitions

Return to our general model, admitting stochastic transitions summarized by \(\tau(s|x, y)\).

5.1 Sufficient Conditions for PAM for all \(\beta\)

Theorem 5  PAM is optimal for all \(\beta\) under either of the following sets of conditions:4

(a) \(\tau(s|\cdot)\) is SBM.
(b) \(Q\) is bi-concave and \(\int_0^z \tau(s|\cdot)ds\) is SBM and bi-concave.
(c) \(Q(x, x)\) and \(\int_0^z \tau(s|x, x)ds\) are concave and \(\int_0^z \tau(s|\cdot)ds\) is SBM.

Proof of (a): Since \(Q\) is SPM, PAM maximizes static output; thus, we need only establish that PAM maximizes continuation values. Since \(\tau(s|\cdot)\) is SBM and \(B(F)(s) = \int \tau(s|\cdot)dF\), PAM solves \(\min_{F \in \Phi(G)} B(F)(s)\) for all \(s\). Since \(V\) increases in FOSD shifts this in turn implies that PAM also solves \(\max_{F \in \Phi(G)} V(B(F))\). Thus, PAM maximizes continuation values.

Proof of part (b): Given \(\int_0^z \tau(s|\cdot)ds\) SBM for all \(z\) PAM solves \(\min_{F \in \Phi(G)} \int_0^z B(F)(s)ds\) for all \(z\). Since \(Q\) and \(\int_0^z \tau(s|\cdot)ds\) are bi-concave, \(V\) is increasing in SOSD shifts by Theorem 1. Together PAM maximizes continuation values. With \(Q\) SPM, PAM maximizes the weighted average of static output and continuation values for all \(\beta\).

Proof of part (c): Assume that for any \(G \succ_{\text{SOSD}} G\) we have \(V(G) \geq V(G)\). We claim this has two implications given our assumptions:

4Parts (b) and (c) are valid when \(V\) rises in the SOSD order. We may use symmetric reasoning to prove results valid when \(V\) increases in the ICX order.
• Step 1: PAM solves the induced Planner’s Problem (1).

• Step 2: $TV(\hat{G}) > TV(G)$,

which together establishes our result.

**Proof of Step 1:** Since $Q$ is SPM, PAM maximizes static payoffs. Since $\int_0^z \tau(s|\cdot)ds$ is SPM, PAM solves $\min_{F \in \Phi(G)} \int_0^z B(F)(s)ds$: that is, PAM minimizes the weight in the tails of the continuation distribution $B(F)$. In turn, this implies that PAM maximizes continuation values given that $V$ increases in SOSD (as assumed). Altogether PAM solves the Planner’s Problem (1) for any distribution $G$.

**Proof of Step 2:** By Step 1, PAM solves (1) for any distribution. But then static output becomes $\int Q(x, x)G(dx)$, which rises in SOSD of $G$ given $Q(x, x)$ is increasing and concave.

Given PAM, the weight in the lower tail of the continuation distribution is

$$\int_0^z B(F)(s)ds = \int_0^z \int \tau(s|x, x)G(dx)ds = \int \int_0^z \tau(s|x, x)ds G(dx),$$

which decreases in SOSD changes since $\int_0^z \tau(s|x, x)ds$ is increasing and concave. □

In each case the given assumptions establish that PAM separately maximizes static payoffs and continuation values. That PAM maximizes static payoffs when $Q$ is SPM is standard, our \( \tau \) assumptions are new.

**Discussion of Transition Assumptions.** Given our discussion of deterministic assumptions, SBM of $\tau$ is clearly a strong assumption. SBM of $\tau(s|\cdot)$ is new. Why is transition SBM critical? The mass in any upper tail of the continuation distribution, $1 - B(F)(s) = \int \int (1 - \tau(s|x, y)) F(dx, dy)$. Thus, by standard reasoning, if $1 - \tau(s|\cdot)$ is SPM for all $s$ (equivalently $\tau(s|\cdot)$ SBM), then PAM maximizes $1 - B(F)(s)$ for all $s$, which implies that PAM maximizes continuation values, since $V$ increases in FOSD shifts.

Since SPM is preserved by addition and the max operator, $f$ SBM implies $\int_0^z \tau(s|x, y)ds$ SPM, but $f$ SPM is not sufficient for $\int_0^z \tau(s|\cdot)ds$ SBM, since SBM is not preserved by the max operator. For example, let $f(x, y) = xy$ then $\int_0^z \tau(s|\cdot)ds$ is not SBM. Thus, although $\int_0^z \tau(s|\cdot)ds$ SBM has some of the flavor of current human capital being complementary in transitions, it is not quite the same.

We can also show that these transition assumptions rule out some typically assumed transitions. For example, human capital follows a martingale and there are two fixed points in the transition process then $\int_0^z \tau(s|\cdot)ds$ SBM cannot obtain. To see this second, assume WLOG the fixed points obtain at $x = 0$ and $x = 1$, and for simplicity assume that $\tau$ is twice differentiable, then we have:

$$\int_0^z \tau(s|1, y)ds = \int_0^z \tau(0, y)ds + \int_0^1 \int_0^z \tau_x(s|x, y)ds dx$$

\footnote{See Anderson and Smith (2007) for a martingale example with fixed points at $x = 0$ and $1$ in which PAM robustly fails to obtain in equilibrium.}
Differentiating both sides with respect to $y$ yields:

\[
\int_0^z \tau_y(s|1,y)ds = \int_0^z \tau_y(0,y)ds + \int_0^1 \int_0^z \tau_{xy}(s|x,y)dsdx \Rightarrow 0 = \int_0^1 \int_0^z \tau_{xy}(s|x,y)dsdx
\]

where the inequalities since given $x = 0$ and $x = 1$ fixed points we have:

\[
\int_0^z \tau(s|0,y)ds = z \forall y > 0 \quad \text{and} \quad \int_0^z \tau(s|1,y)ds = 0 \forall y < 1
\]

so that the derivatives in $y$ evaluated at both $x = 0$ and $x = 1$ are 0. Thus, we cannot have $\int_0^z \tau_{xy}(s|x,y)ds < 0$ for all $(x,y)$.

References


