# Exchanging Good Ideas*

Louise C. Keely

University of Wisconsin

July 3, 2002

*Economics Department, 1180 Observatory Drive, Madison, WI 53706; 608-262-6723; lkeely@ssc.wisc.edu. Two anonymous referees provided helpful suggestions. I thank Philippe Aghion, Paul David, Gilles Duranton, Francisco Ferreira, Markus Haacker, Charles Keely, Rody Manuelli, David Myatt, Henry Overman, Chris Pissarides, Danny Quah, Steve Redding, James Richardson, Tony Venables, Richard Walker, and seminar participants at the World Bank, Oxford, LSE, London Business School, and the University of Wisconsin for helpful comments.

## Abstract

This paper develops a framework to describe how the existence of technological spillovers leads to a distribution of technology clusters. Marshallian spillovers in this paper are assumed to be due to interaction of labor. The dynamics of cluster size, composition, and technology accumulation are characterized. *Journal of Economic Literature* Classification Numbers: D39, O31


Keywords: Technological change, spillovers, coalitions, inequality

*If one man starts a new idea, it is taken up by others and combined with suggestions of their own; and thus it becomes the source of further new ideas.* - Alfred Marshall

Marshall [27] identified the exchange of ideas as a type of externality leading to the localization, or clustering, of economic activity[1]. In Section 1 of this paper, a dynamic model is developed to analyze the formation of groups of skilled workers[2]. Taking Marshall's description of idea exchange as a starting point, the model posits that interaction between heterogenous skilled workers generates spillovers. Patterns of interaction result from clusters of workers. Workers choose with whom they interact, leading to clusters of workers, with technology spillovers occurring within these clusters. Clusters may differ in their levels of technology and technology accumulation. Spillovers take two forms. One is static: the transfer of existing technology. The other is dynamic: the generation of new ideas that lead to new technology. This model can be interpreted as a metaphor of how groups of skilled workers form in a range of settings. The dynamics of cluster size, composition and technology accumulation are explicitly characterized. Two key results of this model are: clusters constitute disjoint subsets along the distribution of heterogeneous workers, and there is a positive long run relationship between the size of clusters and their technology level. A specific example is used to highlight the results.

Implications of the results are considered further in Section 2. It is described how this pa-

[1] The exchange of ideas, alluded to in the quote above, is commonly called technology spillovers and has been used by Henderson [19], Krugman [25], [26] and others to develop explanations for the clustering of economic activity and for differences in income and productivity across geographic space. Feldman [17] provides a general review of the empirical evidence for technology spillovers, both within regions and across countries.

[2] There is a rich theoretical literature on the formation of regions and the evolution of income distribution across those entities (see Alesina and Spolaore [1], Benabou [4], Black and Henderson [6], Casella [12], Durlauf [13], [14], Esteban and Ray [16], Fernandez and Rogerson [18], Rotemberg and Saloner [32], and Tamura [33] ). Nevertheless, these papers do not model the formation of clusters of skilled workers due to dynamic effects of interaction.

per contributes to a theoretical literature on matching across a distribution of individuals with heterogeneous skill. This paper's focus is upon skilled workers particularly, and how clusters of heterogenous workers form as a result of the positive externality they bestow on each other in a cluster. This model can be used to gain insight into the formation of groups of highly-skilled persons such as information technology firms in Silicon Valley, strategy consulting firms like McKinsey, or academic departments. In each of these examples, the production process of each group requires that all members be highly skilled. Short-run and long-run results of this paper's model are compared with related models. This paper provides an alternative interpretation of two types of empirical evidence that have been considered in related work. The first is evidence on trends in the distribution of skill across firms and in income inequality. The second is evidence of a positive relationship between a firm's size and the level of employee skill with a firm. A simple empirical exercise, using data on U.S. economics departments, is used to evaluate one of the model's key results.

# 1 The Model

## 1.1 The Nature of Interaction

As stated above, in this model, technology accumulation occurs through the interaction of skilled workers in clusters[3]. This labor-specific technology is an input into final good production, which will be described below. Cluster formation is described in detail below. Before doing so, the nature of worker interaction and its effects are described.

There are two effects of worker interaction. One is dynamic and is related to the production of new technology; the other is static and involves the catch-up in the level of technology[4].

[3] There are no "unskilled" workers in this model.

[4] These two effects are similar to that used in Jovanovic and Rob [22] in that when workers meet, imitation as

The first is a dynamic effect of interaction on the production of new technology. Interaction between skilled workers in each period produces research ideas; these are used by skilled workers to produce new technology. Time is taken to be discrete. Research ideas are used by the skilled workers to generate new technology according to:

$$\triangle A_{t+1} = A_{t+1} - A_t = \mu n_t^{\gamma_1} \left[ H \left( n_t \right) \right]^{\gamma_2}, \tag{1}$$

$0 < \gamma_1 < 1$, $0 < \gamma_2 \leq 1$, $\mu \in [0,1]$ is a productivity parameter, $n_t$ is the number of workers in the cluster, and $H(n_t)$ is the number of ideas produced, $H'(n_t) > 0$. Examples of functional forms for $H(n_t)$ include pairwise interaction such that $H(n_t) = \frac{n_t(n_t-1)}{2}$ and groupwise interaction[5] such that $H(n_t) = 2^{n_t} - 1$.

This mechanism implicitly assumes that the creation of new technology depends exclusively on worker interaction. The level of previous technology will play a role via the catch-up effect, described next. This isolation of the interaction effect in the dynamic accumulation of technology permits examination of its specific implications for technology cluster formation. One could, for example, include old technology in the technology accumulation function. Doing so would be at the expense of tractability, and its dynamic effect is obvious: it will be a force against convergence in technology levels across clusters. The effect of including the level of existing technology in the technology accumulation function will be further discussed with the results of the paper.

The intuition for (1) is given by example. In an economics department, colleagues talk with each other regularly about research. These conversations, or interactions, lead to new research ideas and potential solutions ($H(n_t)$). It is a reasonable abstraction to suppose the number of these new ideas is dependent on the number of conversations that occur, and the number of conversations is a function of the number of colleagues in the department.

well new invention will occur.

[5]See a related discussion by Weitzman [34] regarding combinations of information.

Turn now to the second effect of interaction: catch-up. This effect is static and occurs upon formation of the cluster, just before the dynamic process of interaction (1). Technology $A$ is ordered on the set of workers. In the model, when skilled workers of different technology levels interact, the ones with lower than the highest technology level in a given cluster must catch up with the highest-technology worker in order to be able to interact with him.

The catch up required for a worker $j$ to interact with the highest technology level worker in the cluster is a function of the gap between their technology levels:

$$C\left(A_t^c, A_{jt}\right) = \left|\left(A_t^c\right)^{\Upsilon} - \left(A_{jt}\right)^{\Upsilon}\right| \leq \omega \tag{2}$$

where $0 < \Upsilon < 1$, $A_{jt}$ is the original technology level of worker $j$ at time $t$, and $A_t^c$ is the technology level of worker $j$ if she joins cluster $c$. This catch-up constraint is asymmetric: if there is a gap in technology levels, the onus of adjustment falls on the workers with the lower levels of technology.[6] If this gap between technology levels is too big, the lower technology workers cannot benefit from being in the cluster and so will choose not to join. The adjustment level may be up to some exogenously-specified constant[7] $\omega > 1$.

When the cluster is first formed, only those with less than the highest level of technology in a cluster gain in the level of technology; this description of how technology levels are exchanged is known as Blackwell ordering [7]. It is described and used by Bhattacharya, Glazer, and Sappington [5] in modelling research joint ventures.

The catch-up constraint (2) is as an assumption regarding the technology of interaction, rather

[6]If one were to change the set up to be a game in which workers were excluded, the equilibrium's main results would be unchanged. For a model with a rule of quality of cluster members, see D. Quah, "Ideas Determining Convergence Clubs", unpublished manuscript, London School of Economics.

[7]This limit could also depend on some function of the payoff from joining the cluster. This more complex formulation does not change the paper's main results.

than an assumption regarding the cost of interaction. The assumption is that interaction technologically cannot take place if the gap in technology between two workers is too large. To express this in terms of cost, one would state that the cost of interaction is zero if (2) holds, and is infinite otherwise.

The constraint is related to the formulation of the rate of technological change by Nelson and Phelps [29]. They assume the rate of technological change is proportional to the gap between the level of technology achieved and a frontier technology. The cost of catching up to a frontier technology is represented by Parente and Prescott [30] and Barro and Sala-i-Martin [2] as increasing in the gap between current and frontier technologies. The constraint (2) is consistent with such an assumption, though, again, the constraint is not itself a cost.

The intuition for the catch-up effect is given by continuing with the example of an economics department. Suppose a faculty member is considering working with one of two PhD students. If this faculty member works with a student, one would not expect his skill as an economist to deteriorate from such interaction. However, it could very well be the case that the student's skill as a research economist will be increased by the interaction. The ability to benefit from the interaction, or to catch up in skill, is dependent upon the student's initial skill level. A student who knows calculus will be better able to catch up and work with the faculty member than one who knows only rudimentary algebra. The latter may be unable ever to catch up with the faculty member, and would not benefit from working with the him or her[8].

The constraint will play a role is determining the pattern of interaction at a point in time. Certain interactions will be prohibited by the constraint. Relaxing the assumption of full catch-up will be discussed in Section 2. In short, the results do not depend on the exact catch-up mechanism

[8]The faculty member, on the other hand, might still benefit from working with the less mathematically-inclined student by talking to him about research in general terms. This possibility is consistent with the model.

specified, so other more complicated mechanisms, with partial or dynamic catch-up, could also be used without changing the main results of the model.

Each worker enters a period with a level of technology accumulated in the previous period. A worker chooses each period to work with others in a cluster, and can work in no more than one cluster. Given the interaction choice, cluster $c$ at time $t$ determines one's current level of technology $A_t^c$ and the number of people with whom one is interacting $N_t^c$.

### 1.2 Final Good Production

There is a single final goods sector that produces output $Q_t$ at time $t$. Output is produced using two factors of production: labor $N_t$ and labor-specific technology $A_t$. All fixed factors are normalized to 1. Technology is distinct from labor and partially excludable: it can be passed from person to person freely, but is attached to labor and can only be passed between workers who interact. Each worker lives one period and has one offspring, ensuring zero population growth. Final good production is described as

$$Q_t = \delta A_t^{\Omega} N_t$$

where $\delta$ is a productivity parameter, $\Omega > 0$ and thus per capita income $q_t$ is

$$q_t = \delta A_t^{\Omega}. \tag{3}$$

The number of workers in the production process will not affect per capita output. Group size effects exist through the knowledge accumulation process only.

### 1.3 Utility

Workers maximize utility over their lifetime and the lifetime of their offspring. Following Durlauf [14] and others, each worker maximizes their offspring's "prospects" - the technology level with which he will enter the next period - and the workers' preference function is

$$U_t^c = u\left(q_t^c\right) + v\left(q_{(t+1)}^c\right) \tag{4}$$

where $c$ is the cluster in which one chooses to work in period $t$. Both $u(\cdot)$ and $v(\cdot)$ are assumed concave, non-decreasing, and continuous functions, where

$$v' \geq 0, v'' \leq 0$$

$$u' \geq 0,\ u'' \leq 0$$

$$u(\infty) = v(\infty) = \infty; \frac{d(u(\infty))}{dq_t^c} = \frac{d(v(\infty))}{dq_{(t+1)}^c} = 0.$$

It is assumed there is no disutility of work. As (4) shows, utility will be determined by the cluster $c$ chosen via the implied $A_t^c$ and $N_t^c$.

### 1.4 Equilibrium Cluster Formation

There are $S$ discrete entities (these may be individuals, firms or regions), indexed by $s$. There is a set of labor of size $LS$ with $L$ workers native to each entity. The initial technology levels are assumed different across entities but initially the same across workers within each entity. Such an assumption does not preclude one from using the model to assess a variety of interaction and spillover situations. Each entity's population of $L$ can individually choose whether to interact with other workers from any other entity including their own. This choice will be described in what follows.

Groups of interacting workers are called clusters. A coalition structure $\pi_t$ consists of a set of $C_t$ clusters, indexed by $c_t = \{1, ..., C_t,\}$ . A coalition structure $\pi_t$ is a partition on the set $LS$ so $\pi_t = \{c_t\}$. The set of all coalition structures is denoted by $\Pi$. Each cluster $c$ contains a nonempty subset of $LS$, consisting of $N_t^c$ persons. Each cluster $c$ at time $t$ has a technology level $A_t^c$ and a level of accumulation $\triangle A_{t+1}^c$. The payoff vector for each worker in the set $LS$ is $U : \Pi \rightarrow \Re$, where $\Re$ is the set of real numbers. In the model described above, given the properties of a cluster, a fixed payoff was assumed, utility, for each skilled worker in each cluster. Thus utility is non-transferable, there is no bargaining over payoff division within a cluster. For any subset $K$ of $LS$, the set of partitions of $K$ is $\Pi_K$ with typical element $\pi_K$.

### 1.4.1 A Description of the Cluster Formation Game

The formation of clusters occurs according to a non-cooperative game. The structure of the non-cooperative game utilized is similar to that developed by Bloch [8]. An informal description of the game is provided here, and the reader may refer to Bloch's paper for complete details. To be clear, this paper uses Bloch's coalition game to solve for an equilibrium in each period. The model of technology spillovers for which an equilibrium is solved and the dynamic analysis of the evolution of technology spillovers and technology clusters are specific to this paper.

There is a rule of order $\rho$ for $LS$, used to determine the order of moves in the sequential game of cluster formation. The game is called $\Gamma(U, \rho)$ because cluster formation may depend on

- the specification of preferences, which derives from any element of $\Pi$ and the implied $(N_t^c, A_t^c)$ given a set $LS$ and $\{A_{jt}\}_{j \in LS}$, and
- the rule of order $\rho$ for a given set $LS$.

Clusters are formed sequentially in each period $t$ according to the game $\Gamma(U, \rho)$. Each player knows the technology levels of the set $LS$, $\{A_{jt}\}_{j \in LS}$. The first skilled worker according to $\rho$ proposes the formation of a cluster $T$ to which she will belong. Each member of $T$ responds to the proposal in the order given by $\rho$. If one player rejects the proposal, this player immediately makes a counteroffer and proposes a cluster $T'$ to which she will belong. On the other hand, if all members accept, the cluster is formed. All members of $T$ would then form the cluster and leave the game, and the first skilled worker in $LS\backslash T$ according to $\rho$ makes a proposal. Note that once a cluster is formed, the game is then played only by the remaining players. Once a player joins a cluster, she cannot leave or propose to change the cluster in that period.[9]

A skilled worker $j$ is active, at time $t$ and stage $r$ with history $h_{tr}$, if it is her turn to accept/reject or propose after the history $h_{tr}$. The set of histories at period $t$ at which a skilled worker $j$ is active is $H_{jt}$. A strategy $\sigma_{jtr}$ for a skilled worker $j$ is a mapping from $H_{jt}$ to her set of actions, that is

$$\sigma_j(h_{tr}) \in \{Yes, No\} \text{ if } \hat{T}(h_{tr}) \neq \emptyset$$

$$\sigma_j(h_{tr}) \in \left\{T \subset LS\backslash\hat{K}(h_{tr}), j \in T\right\} \text{ if } \hat{T}(h_{tr}) = \emptyset$$

where $\hat{K}(h_{tr})$ is the set of skilled workers who will have left the game given history $h_{tr}$, and $LS\backslash\hat{K}(h_{tr})$ is the set of workers remaining in the game.

Due to the structure of $U$, the only utility-relevant parts of the game's history are the set $K$ of skilled workers who have left the game, the implied set $LS\backslash K$ and their current technology levels, and the current proposed cluster formation $T \in \{LS\backslash K\}$. A worker will choose a strategy $\sigma_{it}$, and that strategy is stationary if it only depends on the state[10] $\left(K, LS\backslash K, T, \{A_{jt}\}_{j\in LS\backslash K}\right)$. In particular, a stationary strategy does not depend on the payoffs of those who have already left the game, nor on the history of rejections. A strategy profile $\sigma_t = \{\sigma_{jt}\}_{j\in LS\backslash K}$ determines the outcome $\pi(\sigma_t)$.

[9] Relatedly, no bargaining within a cluster can occur due to the non-cooperative framework and the absence of an enforcement mechanism. For a model of coalition formation with intra-coalition bargaining, see Ray and Vohra [31].

[10] This definition of stationarity specifically includes the identifying characteristics of the players. The exact form of heterogeneity is not specified in the more abstract models of Bloch [8] and Moldavanu and Winter [28]. However, none of the results depending on those derived in Bloch [8] and Moldavanu and Winter [28] are altered since their models do not exclude the possibility of heterogeneity among players.

### 1.4.2 Equilibrium Definition and Properties

The concept of equilibrium in this model is now clarified and its features described. All players are part of some cluster, even if it is a one worker cluster which is the outcome of joining no other multi-worker cluster. The payoff for a player $j$ is given by $U_j(\pi(\sigma_t)) = U(N_t^c, A_t^c,\ j \in c)\ \forall j \in LS$.

**Definition 1** *A subgame perfect equilibrium $\sigma_t^*$ is a strategy profile such that $\forall j \in LS$, $\forall h_{tr} \in H_t^j$, $\forall \sigma_{jt}$, $U_j\left(\pi\left(\sigma_{jt}^*, \sigma_{-jt}^*\right)\right) \geq U_j\left(\pi\left(\sigma_{jt}, \sigma_{-jt}^*\right)\right)$.*

**Definition 2** *A stationary perfect equilibrium $\sigma_t^*$ is a subgame perfect equilibrium where $\forall j \in LS$, $\sigma_{jt}^*$ is a stationary strategy.*

**Definition 3** *A perfectly stratified equilibrium is one in which coalitions are formed that partition into disjoint subsets the distribution of technology levels $\{A_{jt}, j \in LS\}$.*

**Definition 4** *An order of play independent equilibrium is one in which, for a strategy profile $\sigma_t$ that is a subgame perfect equilibrium for the game $\Gamma(U, \rho)$, the payoff for each player $j$ is given by $U_j(\pi(\sigma_t)) = U(N_t^c, A_t^c,\ j \in c)$ and does not depend on $\rho$ (Moldavanu and Winter [28]).*

**Definition 5** *Consider $T$ time periods over which the game of this model is played in each period, with each time period indexed by $t \in [0, T]$. The limiting distribution is the distribution of the technology level of workers that results as $T \to \infty$.*

**Proposition 1** *A stationary perfect equilibrium $\sigma_t^*$ exists with the following characteristics:*

1. *At each period $t$, the clusters are perfectly stratified;*

2. *At each period $t$, a subgame perfect equilibrium is independent of the order of play $\rho$ and therefore coincides with the core of the game;*

3. *In the limiting distribution, ordering clusters by technology level or by size is equivalent.*

The proof of the proposition is in the Appendix. The logic of the proof is summarized here. At each point in time in equilibrium clusters form in order of technology level, as indicated in Figure 1. To show perfect stratification, it is first established that utility is increasing, all else equal, in cluster size. Therefore, the existence of more than one cluster is due to the catch-up constraint (2) , which makes some interactions infeasible. The proof thus proceeds by showing that the combined sets of feasible interactions for any two workers with different initial technology levels constitutes a superset of the set of feasible interactions for any worker with a skill level between those two levels. The choice set of feasible clusters for the workers with the highest and lowest skill levels is at least as large as that for all the workers in between. If the two workers with the highest and lowest technology levels propose or accept to interact in the same cluster in equilibrium, so will all workers with technology levels in-between.. This argument is applicable to any subset of the entire technology distribution, and thus to the entire distribution. Perfect stratification follows immediately.

To establish coincidence of an equilibrium with the core, the irrelevance of the order of play is noted. A cluster will emerge in equilibrium if and only if, for any member of that cluster, there is no other potential cluster at any stage of the game with overlapping membership that is utility-superior, feasible, and would be accepted by all others in that potential cluster. The order of play thus does not affect subgame perfect equilibrium strategy of each worker. Order independence is implied. Stationarity in an equilibrium , so that only payoff-relevant elements of the model affect equilibrium strategies, results from a degree of unanimity in preferences and a lack of circularity in the players' strategies. Using a result from Moldavanu and Winter [28], it is possible to conclude that this equilibrium coincides with the core of the game. The non-cooperative coalition formation game does not yield a different outcome than a cooperative game would.

For the intuition behind the long run equivalence between ordering clusters by size and technology level, suppose there are two clusters of different size at time $t$. Members of the larger cluster will have a higher $\triangle A_{t+1}$ than those of the smaller cluster. If members of the larger cluster also have higher initial technology level $A_t$ then the gap in technology levels between the clusters will increase. Since two clusters will exist only due to the catch-up constraint (2), then the constraint can continue to bind and the cluster formation persist. The technology level and size of clusters will be ordered equivalently. If, however, the members of the larger cluster have the lower initial technology level, then the gap in technology levels between the clusters will decrease. As long as the catch-up constraint binds, then this cluster formation will persist and the gap will continue to narrow. Eventually, the two clusters will merge and, trivially, the technology level and size of clusters will be ordered equivalently.

If $\triangle A_{t+1}$ were to be increasing not only in interaction but somehow in the initial levels of technology of the cluster members, then the long run result of equivalent ordering of clusters by size and technology level would no longer generally hold. The long run results would depend delicately on the relative importance of interaction and original technology level in the function $\triangle A_{t+1}$. Isolating these two effects is therefore useful. It is obvious that the inclusion of $A_t$ in $\triangle A_{t+1}$ would be a force for divergence of clusters after the initial period. This would work against the catch-up process of larger lower technology clusters, and would reinforce the divergence between larger higher technology clusters and their followers. Not including $A_t$ in $\triangle A_{t+1}$ may sacrifice some generality. The gain is in the ability to isolate the effect of interaction on the dynamics of cluster formation and to derive clear results regarding that effect. These results are not a priori obvious.

### 1.5 An Example

A simple example describing the evolution of cluster formation should assist in making concrete the results of the model. Consider a uniform initial distribution of the level of technology $A$ for $N$ workers on the range $[1, N]$. Workers are indexed by their initial technology level. The constraint on interaction is reproduced here for convenience:

$$C\left(A_t^c, A_{jt}\right) = \left|\left(A_t^c\right)^{\Upsilon} - \left(A_{jt}\right)^{\Upsilon}\right| \leq \omega.$$

The accumulation of technology for a worker in a given cluster $c$ of size $n$ is, again, given by:

$$\triangle A_{t+1}^c = A_{t+1} - A_t = \mu n_t^{\gamma_1}\left[H\left(n_t\right)\right]^{\gamma_2}.$$

The utility derived from interaction in a cluster $c$ can be written as a function of the technology level and accumulation of that cluster:

$$U_t^c = u\left(\delta\left(A_t^c\right)^{\Omega}\right) + v\left(\delta\left(A_t^c + \triangle A_{t+1}^c\right)^{\Omega}\right)$$

The following parametric values are used for the example:

$$\begin{aligned} \Upsilon &= 0.5 \\ \omega &= 1.0 \\ N &= 64 \\ H\left(n_t\right) &= n_t \\ \gamma_1 + \gamma_2 &= 0.5. \end{aligned}$$

These values are consistent with the assumptions made above and were chosen to be reasonable as well as to provide a simple but illuminating illustration of the model's dynamics in the short and long run. As $\Upsilon$ increases and $\omega$ decreases, the constraint becomes more restrictive. The choice to have $n_t$ enter with decreasing returns into the technology accumulation function is made to provide more interesting short run dynamics.

The evolution of cluster formation and the technology distribution is summarized in Figure 2 below, while Figure 3's table provides details of the exercise. In the initial period in which clusters are formed, one may determine the composition of the clusters by starting with the highest technology level worker $N = 64$ and determine who is the lowest technology level worker that can interact with worker $N$. That is worker 49. A cluster composed of 16 workers is the largest feasible cluster that each of them can join, given the catch-up constraint, and this cluster will also obviously yield the highest technology level. Thus, the highest cluster that will form in equilibrium is called $C1$ and consists of 16 workers over the range $[49, 64]$. The same exercise is now repeated, starting with worker 48. The lowest technology level worker who can interact with worker 48 is worker 36. The largest feasible cluster that these workers can interact in is size 13, and utility maximization is achieved by interacting with the highest technology level possible, i.e. 48. Thus, the second highest cluster that will form in equilibrium is called $C2$, consisting of 13 workers over the range $[36, 48]$. Repeating this exercise down the initial technology distribution, cluster $C3$ is $[25, 35]$; cluster $C4$ is $[16, 24]$, cluster $C5$ is $[9, 15]$, cluster $C6$ is $[4, 8]$ and, finally, cluster $C7$ is $[1, 3]$. The fact that the clusters are initially decreasing in size is an artefact of the uniform technology distribution in this example and is not a general property of the model. In this example, the irrelevance of the order of play is clear, allowing one to determine the equilibrium cluster formation without directly appealing to the game theoretic framework.

In each cluster, all workers immediately catch-up to the initial technology level of the highest-technology worker in the cluster. Thus, the distribution goes from being uniform to consisting of

7 discrete points: 64, 48, 35, 24, 15, 8 and 3 in the initial period $T = 0$. Each of these clusters will accumulate technology according to its size. In period $T = 1$ one can then again use the constraint to evaluate whether workers from $C1$ and $C2$ can interact, whether workers from $C2$ and $C3$ can interact, and so on. Once these workers have interacted, they will always interact, so the only issue is whether clusters will merge to form larger clusters. To do so, the difference in technology levels must be sufficiently small. It turns out that the initial clusters $C2$ and $C3$; $C3$ and $C4$, and $C4$ and $C5$ can interact. Since the workers in $C3$ will choose to interact with workers in $C2$ to form a cluster of size 24, it will be the case that $C4$ and $C5$ workers will also interact to form a cluster of size 16. Clusters $C1$, $C6$, and $C7$ have the same composition.

The second highest cluster, made up of $C2$ and $C3$ workers, is now larger than the highest cluster $C1$. This illustrates that the size distribution need not remain the same over the short run, and that in the short run the technology level and size of clusters may not be equivalently ordered. Cluster $C2 + C3$ therefore has a higher level of technology accumulation from periods 1 to 2, and in the next period $T = 2$, the cluster $C2 + C3$ is able to interact with $C1$ and this interaction will be chosen in equilibrium in the next period. Cluster $C6$ cannot interact with cluster $C4 + C5$, but clusters $C6$ and $C7$ can interact[11].

In the third period $T = 3$ the long run stratified pattern of clusters has formed, although the level of technology for each cluster will continue to change over time. The highest cluster, consisting of the original clusters $C1 + C2 + C3$ is size 40 and is the largest cluster. The middle cluster, consisting of the original clusters $C4 + C5$ is size 16, and the lowest cluster, consisting of the original clusters $C6 + C7$ is size 8. The positive size-technology relationship is clear. There is not divergence of clusters in the strict sense, as the growth rate of technology in the larger

[11] The fact that clusters $C6$ and $C7$ can interact in $T = 2$ but could not in $T = 1$ illustrates that the constraint may become weaker as technology accumulates since $\omega$ is unchanging over time.

clusters is smaller than in the lower clusters. A measure of inequality, the coefficient of variation of technology, is decreasing asymptotically toward a constant although the variance itself is increasing. The interaction constraint continues to bind over time. In fact, the gap in technology level between agents in different clusters is increasing over time, as the cluster formation game and interaction occurs in each period.

# 2 Discussion

This paper develops a framework for understanding how endogenous technology spillovers can lead to a distribution of technology clusters. Technology spillovers in this paper are taken to be due to interaction of labor, as in Marshall [27]. Features of the distribution at a point in time and its evolution are identified. Examining the distribution over time, as in the example above, is interesting because the long run distribution patterns of size and technology across clusters may differ from of those of the short run. The fact that model settles down to a long run stable cluster pattern allows for the statement of general long run results and a simple empirical test of the model.

## 2.1 Dynamics of Inequality and Skill Segregation

Other models have been used to examine the causes and consequences of skill segregation. These models looking at related issues have either used static models or focused on steady state results. Understanding why skill segregation occurs can be important from a welfare point of view. Even in the simple setting of this paper's model, if the social utility function is simply the sum of utility, integration of all workers is the utility-maximizing outcome, and so a decentralized equilibrium would generally not be welfare-maximizing. Burdett and Coles [10] study a matching model, often used in models of employment, in order to analyze marriage choices across a distribution of heterogenous agents with non-transferable utility, and continual replacement of the population as people marry and exit. They show that one can divide the individuals into classes by attractiveness, and that marriages will take place only within those classes in equilibrium. They characterize the steady state equilibria, and use examples to further highlight different equilibrium outcomes.

Both Kremer [23] and Kremer and Maskin [24] use matching models in order to examine the distribution of employee skill across employers. Kremer [23] has a result of perfect sorting by skill across firms. Kremer uses the model to provide explanation for several cross-sectional facts such as variation in product quality across firms, and the extremely large variation in income across countries. Kremer and Maskin [24] present evidence that there has been increasing segregation of workers by skill across firms. They develop a model to understand this trend. Their model is static, and so they examine one-time changes in the economy. They show that either exogenous increases in the mean skill level or in the dispersion of the skill distribution can produce an increase in segregation and inequality of skill across firms. Their model contains a broad range of skill groups, and so also suggests explanations for phenomena such as how decreases in wages of the low skilled can coincide with increases in wages of the high skilled.

In this paper's model, as in those just described, workers are segregated by skill across firms and the variance[12] in skill is increasing in the long run. The model is dynamic so that the distribution of technology and the choice of cluster both change endogenously to give the result. The game

[12] The increase in variance is consistent with the increase of dispersion in Kremer and Maskin [24], which is essentially a measure of the distance between the highest and lowest skilled workers. If inequality is measured in other ways, such as the coefficient of variation or using the Gini coefficient, it is not clearly the case that in the long run inequality would be rising, although certainly in the short run inequality could be rising. However, it is not obvious that the recent increase in income inequality in the U.S. is a long run phenomenon.

theoretic framework produces simple rules regarding cluster formation in each period's equilibrium, allowing for precise characterization of the dynamics in examples such as that of the previous section. This model suggests not only why interaction of skilled workers leads to skill segregation and inequality persistence in the long run, but that these stratification patterns may not persist in the short run. The segregation of workers by skill - consistent with all three models described above - is a short and long run phenomenon in the model, but increasing skill dispersion may not occur in the short run alongside segregation. Thus, if skill dispersion were to decrease[13], skill segregation would not reverse in this model as it could in Kremer and Maskin [24].

The framework of this paper allows only for temporary inequality in skill within a cluster because of immediate catch-up with the highest technology level worker in the cluster. This assumption is maintained for simplicity of exposition and would be straightforward to relax, to allow for only partial catch-up within each cluster. Inequality would be present within clusters, although there would be convergence of skills within clusters with continued interaction over time. This result is consistent with the effect of technological change on the distribution of skill within clusters in Kremer and Maskin [24]: firms become more homogenous. However, the reasoning is quite different. In Kremer and Maskin's model, exogenous increases in the mean skill level leads to the composition of firms changing. In this model, the increasing homogeneity of skill occurs because of interaction and thus does not necessarily imply a change in cluster composition.

[13]Income inequality has risen since the late 1960's, although it appears to have plateaued, or even declined, in the last few years (Jones and Weinburg [21]). Without making too much of just a few years of data, it is still appropriate to note that the rise in inequality that the U.S. has experienced may not continue forever. The question of whether the link between inequality and skill segregation is symmetric in the direction of change in inequality is thus important to consider.

### 2.2 The Size-Skill Relationship

A main result of this paper's analysis is that there is a long run positive relationship between the size of a cluster and its technology level. The term technology is used here, although one could equivalently call it *skill* in the context of the model since technology is worker-specific and the result of worker interaction. Therefore, this model's findings lend support to the hypothesis that the positive size-wage relationship across firms is due to higher skill workers being employed by larger firms[14]. This hypothesis has substantial empirical support (see Bayard and Troske [3], Brown and Medoff [9], and Idson and Oi [20]). A model that has been offered as an explanation for this evidence is employer-employee matching such as that in Burdett and Coles [10], so that better employers hire better employees, and larger firms offer a better workplace environment than smaller firms, all else equal. Most of the matching literature focuses on a fixed size 2 cluster (see Burdett and Coles [10]). However, Kremer [23] matches each worker with one task in a setting where several tasks may together constitute a single production process and thus a single firm. When more tasks are linked together, the complexity of each individual task increases. The marginal product of each worker is increasing in the skill of other workers at the same firm. Thus, in equilibrium, firms that undertake higher complexity production processes will hire higher skilled workers. By assuming that larger firms have more complex production processes, Kremer links skill segregation with the size-skill hypothesis. This explanation relies on a worker-task matching framework, as well as the assumption that large firms produce more complex goods.

This paper offers an alternative, though complementary, explanation for a long-run positive relationship between the size of a firm and the skill of workers in the firm. The alternative contains two ingredients. All else equal, workers themselves prefer to be with other workers of higher skill.

[14] An interesting, though only loosely related, empirical fact is documented by Black and Henderson [6]. They examine the distribution of US cities and document that as human accumulation grows, so does city size.

In addition, skill accumulation of each worker is increasing in the number of other workers with whom they interact. The dynamic effect of interaction between workers is necessary for the skill-size relationship in the long run.

Because this paper's model does not appeal to a matching framework, if the model has validity as a model of interaction between skilled workers, its empirical relevance should go beyond that of the firm employer-employee setting. For example, it should be relevant for understanding the composition and distribution of academic departments. In academic departments there is no individual boss or employer, and the basis for hiring is the quality of research and interaction expected from a prospective employee. The model predicts that department size should be positively correlated with department reputation in research skill. As an exercise to test this prediction, it is possible to look at the distribution of 50 economics departments in the United States whose faculty are active in research. These are the departments who are listed in the Dusansky-Vernon [15] ranking of economics departments. The data are presented in the table in Figure 4. It was found that:

1. The departments do differ statistically significantly in ranking, when a cardinal ranking is used. This cardinal ranking is determined by the adjusted number of pages published per faculty, where the adjustments are for journal quality. A Chi-squared test is used. (Figure 5)

2. The departments do differ statistically significantly in size, both in absolute level and in faculty per undergraduate student. A Chi-squared test is used. (Figure 5)

3. There is a statistically significant positive correlation between the cardinal rank of a department and the size of a department, controlling for possible exogenous funding variations proxied by the size and public/private status of the school. (Figure 6)

The results are consistent with this paper's model of worker interaction, and with the empirical literature that posits that the positive relationship between firm size and wages is due to higher skilled workers being employed by larger firms. Consistency with the results of the model of Section 1 indicates that economics departments are limited in size because technology spillovers are more difficult the greater is the distance in quality between two researchers. Further, the model of Section 1 suggests that larger economics departments produce more spillovers between the faculty in the department.

# 3 Appendix

## 3.1 Proof of Proposition 1

**A stationary perfect equilibrium $\sigma_t^*$ exists with the following characteristics:**

The existence of a subgame perfect equilibrium at time $t$ is established by Bloch [8]: Proposition 2.4 and Corollary 2.5. Stationarity is established below once order of play irrelevance is shown.

**At each period $t$, the clusters are perfectly stratified;**

From the construction of technological accumulation within a cluster (1), (3), and utility (4), it is straightforward to show that

$$\frac{\partial U_t^c}{\partial A_t^c} > 0; \frac{\partial U_t^c}{\partial N_t^c} > 0.$$

Absent the catch-up constraint (2), a single cluster would form in equilibrium. Therefore, if there exists $C_t > 1$ clusters at time $t$ in equilibrium, then it must that the case, for any two clusters $c_1, c_2 \in C_t$ that at least one member of $c_1$ or $c_2$ is unable to interact with at least one member of the other cluster due to the catch-up constraint.

To establish the feasibility and then the optimality of perfect stratification, a period of interaction $t$ is examined. Consider two (groups of) workers, $s_1$ and $s_2$, with technology levels $A_{s_1t} < A_{s_2t}$. Next, consider the set of $M \geq 1$ workers, indexed by $k$, such that $A_{s_1t} \leq A_{mt} \leq A_{s_2t}$, $m \in M$. It is shown that if workers of technology level $A_{s_1t}$ accept or make a feasible proposal to interact with workers of technology level $A_{s_2t}$, or vice versa, then workers of technology level $A_{s_2t}$ will also be included in and accept that proposal in equilibrium.

a. If interaction is feasible between the $s_1$ and $s_2$ workers, then it must be feasible between $s_1$ or $s_2$ and any of the $m \in M$ workers. That is, if:

$$\left|(A_{s_2t})^{\Upsilon} - (A_{s_1t})^{\Upsilon}\right| \leq \omega \tag{5}$$

then since $A_{s_1t} < A_{mt} < A_{s_2t}$,

$$\left|(A_{s_2t})^{\Upsilon} - (A_{mt})^{\Upsilon}\right| \leq \omega$$

and

$$\left|(A_{mt})^{\Upsilon} - (A_{s_1t})^{\Upsilon}\right| \leq \omega.$$

b. If interaction is feasible between the $s_1$ and $s_2$ workers, that is (5) holds, then for any $\tilde{A}_t \in [A_{s_1t}, A_{s_2t}]$ :

$$\left|\left(\tilde{A}_t\right)^{\Upsilon} - (A_{mt})^{\Upsilon}\right| \leq \omega,$$

$$\left|\left(\tilde{A}_t\right)^{\Upsilon} - (A_{s_2t})^{\Upsilon}\right| \leq \omega,$$

and

$$\left|\left(\tilde{A}_t\right)^{\Upsilon} - (A_{s_1t})^{\Upsilon}\right| \leq \omega.$$

c. For any $\tilde{A}_t \geq A_{s_2t}$

$$\left|\left(\tilde{A}_t\right)^{\Upsilon} - (A_{mt})^{\Upsilon}\right| \leq \omega$$

only if

$$\left|\left(\tilde{A}_t\right)^{\Upsilon} - (A_{s_2t})^{\Upsilon}\right| \leq \omega.$$

d. For any $\tilde{A}_t \leq A_{s_1t}$

$$\left|(A_{mt})^{\Upsilon} - \left(\tilde{A}_t\right)^{\Upsilon}\right| \leq \omega$$

only if

$$\left|(A_{s_1t})^{\Upsilon}-\left(\tilde{A}_t\right)^{\Upsilon}\right|\leq\omega.$$

By a-d, the combination of the two sets of technology levels that are feasible for interaction by the workers $s_1$ and $s_2$ constitutes a superset of the set of technology levels that are feasible for interaction by any $m \in M$ worker such that $A_{s_1t} \leq A_{mt} \leq A_{s_2t}$, $m \in M$.

Thus, if any two (groups of) workers $s_1$ and $s_2$ propose or accept to interact in the same cluster in equilibrium, then so will each worker $m \in M$. Because utility is increasing in cluster size, any equilibrium that includes workers $s_1$ and $s_2$ will also include all workers $m \in M$.

Perfect stratification in equilibrium follows immediately by extending the above argument to any part of the distribution of workers, including the entire distribution, at any time $t$.

**At each period $t$, a subgame perfect equilibrium is independent of the order of play $\rho$ and therefore coincides with the core of the game;**

To ignore an uninteresting case of indifference between potential clusters, assume:

**Assumption 1** Suppose a worker $i$ is indifferent in terms of utility outcomes between $R$ alternative potential clusters in which it would be a member of each, where the clusters are denoted as subsets $\{T_1, ..., T_R\} \in \Pi_{LS \backslash K}$. Then worker $i$ will randomly choose which cluster to propose or accept when it is her turn in the order of play $\rho$.

Suppose a cluster of composition $T \in \Pi_{LS\backslash K}$ results in equilibrium at stage $r$ of time $t$. Then this cluster must be feasible, under the catch-up constraint (2) for all members of $T$. By the last section of the proof, it is also known that this cluster $T$ will constitute a disjoint subset of the partition of the distribution of workers' technologies. By construction of the preferences, catch-up constraint, and the cluster formation game, at any stage $r$ with $LS\backslash K$ the set of players remaining (and $K = \emptyset$ is possible):

a.$T$ will have been proposed and accepted by all members *only if* $\nexists$ cluster $T' \in \Pi_{LS}$ where $T'$ would have the following characteristics:

- there is at least partial overlap in worker membership between $T$ and $T'$
- $U\left(N_t^T, A_t^T,\ j \in T\right) < U\left(N_t^{T'}, A_t^{T'},\ j \in T'\right)$ for any worker $j$
- the cluster $T'$ is itself feasible for all potential members
- $T'$ would be accepted by each member $j \in T'$, $j \notin T$.

b. *If* $\nexists$ cluster $T' \in \Pi_{LS}$ with the characteristics just described, $T$ would be proposed and accepted at some stage $r$, $T \in \Pi_{LS \backslash K}$.

The arguments of points a-b are order of play irrelevant. The cluster $T$ would be proposed and accepted at any stage $r$ only if no other cluster $T''$ would preempt $T$ at an earlier stage, and only if no other cluster $T'$ possible amongst the remaining players at stage $r$ provides a utility-superior and feasible alternative to the overlapping memberships. Conversely, if $T$ is to be accepted at stage $r$, it will not be preempted by another cluster $T''$ this is either utility-inferior or not feasible to all overlapping members.

It is implied by these points that each subgame perfect equilibrium strategy of a player $i$, $\sigma_{it}^*$, holding fixed others' strategies $\sigma_{-it}^*$, is order irrelevant. In other words, the strategy is the same regardless of the order of play. This implies order independence, which is a weaker condition. For a given subgame perfect equilibrium strategy of a player $i$, that player's payoff does not depend on the order of play $\rho$. This is order independence.

Order irrelevance for the subgame perfect equilibrium strategy also implies that the strategy chosen depends only on the payoff relevant state, and not on the payoffs of those who have left the game or the history of rejections. Thus, a subgame perfect equilibrium strategy is stationary.

The coincidence between an order independent stationary equilibrium and the core follows directly from Moldavanu and Winter [28] Propositions A and B and Corollary 2. They show that a stationary perfect equilibrium that is order independent coincides with the core of the coalitional game.

**In the limiting distribution, ordering clusters by technology level or by size is equivalent.**

It remains to show that in the limit, cluster technology levels are larger, the larger is the cluster size.

It is noted that once a group of workers interact in the same cluster in one period, they will do so in all future periods. This follows directly from the facts that

- For any two workers $i$ and $j$ who interacted in period $t$ and who consider interacting in period $(t+1)$, such interaction is feasible:

$$C\left(A_t^c + \triangle A_{t+1}^c, A_t^c + \triangle A_{t+1}^c\right) = 0,$$

- in period $t+1$, for some worker $k$, if interaction with the worker $i$, referred to in the previous point, is feasible, it must also be feasible with the worker $j$,
- and $\frac{\partial U_{t+1}^c}{\partial N_{t+1}^c} > 0 \; \forall t, c.$

Thus, any equilibrium cluster in period $t+1$ that includes the worker $i$ must also include the worker $j$. This argument extends directly to a larger group of workers who interacted in time $t$, and also can be repeated in all periods $t+s$, $s \geq 1$. Therefore, for any worker $i$, the cluster $c$ to which she belongs at time $t$, of size $N_t^c$, and the cluster $c'$ to which she belongs at time $t'$, $t' > t$, it will be the case that $N_t^c \leq N_{t'}^{c'}$. In other words, the average size of clusters will be non-decreasing over time.

Next, suppose there are two clusters $c$ and $d$. If $A_t^c > A_t^d$, then $\triangle A_{t+1}^c < \triangle A_{t+1}^d$ if and only if $N_t^c < N_t^d$. If $N_t^c < N_t^d$, then there exists some time period $\tilde{t} > t$ such that either cluster $d$ will overtake cluster $c$: $A_{\tilde{t}}^c < A_{\tilde{t}}^d$ (if $N_t^d$ is sufficiently large relative to $N_t^c$), or where clusters $c$ and $d$ will join in a single cluster.

On the other hand, if $A_t^c > A_t^d$ and $N_t^c > N_t^d$, then these two clusters would not be a single cluster only because $\left|(A_t^c)^\Upsilon - \left(A_t^d\right)^\Upsilon\right| > \omega$. However, if $\left|(A_t^c)^\Upsilon - \left(A_t^d\right)^\Upsilon\right| > \omega$ and $N_t^c$ is sufficiently larger than $N_t^d$, $N_t^c >> N_t^d$, then $\left|(A_t^c)^\Upsilon - \left(A_t^d\right)^\Upsilon\right|$ will be non-decreasing over time and so $\left|\left(A_{\tilde{t}}^c\right)^\Upsilon - \left(A_{\tilde{t}}^d\right)^\Upsilon\right| > \omega \ \forall \tilde{t} > t$. The two clusters will never merge if and only if $N_t^c$ is sufficiently larger than $N_t^d$.

In sum, the two clusters will either merge into one cluster, or the two clusters will remain separate indefinitely and be equivalently ordered by size and technology level.

It is conceptually straightforward to apply this argument to $C_t > 2$ clusters.

Therefore, as $t \to \infty$, equilibrium clusters must be equivalently ordered by technology or size.■

**References**


1. A. Alesina and E. Spolaore, On the number and size of nations, *Quart. J. Econ.,* **112**(1997), 27-56.

2. R. J. Barro and X. Sala-i-Martin, Technology diffusion, convergence and growth, *J. Econ. Growth,* **2**(1997), 1-26.

3. K. Bayard and K.R. Troske, Examining the employer-size wage premium in the manufacturing, retail trade, and service industries using employer-employee matched data, *Amer. Econ. Rev.,* **89**(1999), 99-103.

4. R. Benabou, Workings of a city: Location, education and production, *Quart. J. Econ.,* **108**(1993), 619-652.

5. S. Bhattacharya, J. Glazer, and D.E.M. Sappington, Licensing and the sharing of knowledge in research joint ventures, *J. Econ. Theory,* **56**(1992), 43-69.

6. D. Black and V. Henderson, Urban growth, *J. Polit. Economy,* **107**(1999), 252-284.

7. D. Blackwell, Comparison of experiments, in "Proceedings of the Second Berkeley Symposium on Probability and Mathematical Statistics" (J. Neyman, Ed.), pp. 93-102, University of California Press, Berkeley and Los Angeles, 1951.

8. F. Bloch, Sequential formation of coalitions in games with externalities and fixed payoff division, *Games Econ. Behav.,* **14**(1996), 90-123.

9. C. Brown and J. Medoff, The employer size-wage effect, *J. Polit. Economy,* **97**(1989), 1027-1059.

10. K. Burdett and M.G. Coles, Marriage and class, *Quart. J. Econ.,* **112**(1997), 141-168.

11. K. Burdett and M. G. Coles, Long-term partnership formation: Marriage and employment, *Econ. J.,* **109**(1999), F307-F334.

12. A. Cassella, The role of market size in the formation of jurisdictions, *Rev. Econ. Stud.,* **68**(2001), 83-108.

13. S.N. Durlauf, Spillovers, stratification and inequality, *Eur. Econ. Rev.,* **38**(1994), 836-845.

14. S.N. Durlauf, A theory of persistent income inequality, *J. Econ. Growth,* **1**(1996), 836-845.

15. R. Dusansky and C.J. Vernon, Rankings of U.S. economics departments, *J. Econ. Perspect.,* **12**(1998), 157-170.

16. J.M. Esteban and D. Ray, On the measurement of polarization, *Econometrica,* **62**(1994), 819-851.

17. M.P. Feldman, The new economics of innovation, spillovers and agglomeration: A review of empirical evidence, *Econ. Innov. New Tech.,* **8**(1999), 5-25.

18. R. Fernandez and R. Rogerson, Income distribution, communities, and the quality of public education, *Quart. J. Econ.,* **111**(1996), 135-164.

19. J.V. Henderson, The size and types of cities, *Amer. Econ. Rev.,* **64**(1974), 640-656.

20. T.L. Idson and W.Y. Oi, Workers are more productive in large firms, *Amer. Econ. Rev.,* **89**(1999), 104-108.

21. A.F. Jones and D.H. Weinberg, The changing shape of the nation's income distribution, U.S. Census Bureau Report P60-204, Washington D.C., 2000.

22. B. Jovanovic and R. Rob, The growth and diffusion of knowledge, *Rev. Econ. Stud.,* **56**(1989), 569-582.

23. M. Kremer, The o-ring theory of economic development, *Quart. J. Econ.,* **108**(1993), 551-575.

24. M. Kremer and E. Maskin, Wage inequality and segregation by skill, NBER Working Paper 5718, 1995.

25. P. Krugman, "Geography and Trade," MIT Press, Cambridge, Massachusetts, 1991.

26. P. Krugman, Innovation and agglomeration: Two parables suggested by city-size distributions, *Japan World Econ.,* **7**(1995), 371-390.

27. A. Marshall, "Principles of Economics," 8th edition, Macmillan, New York, 1920.

28. B. Moldavanu and E. Winter, Order independent equilibria, *Games Econ. Behav.,* **9**(1995), 21-34.

29. R.R. Nelson and E.S. Phelps, Investment in humans, technological diffusion, and economic growth, *Amer. Econ. Rev.,* **56**(1966), 69-75.

30. S. Parente and E. Prescott, Barriers to technology adoptioon and development, *J. Polit. Economy,* **102**(1994), 298-321.

31. D. Ray and R. Vohra, Coalitional power and public goods, *J. Polit. Economy,* **109**(2001), 1355-1384.

32. J. Rotemberg and G. Saloner, Competition and human capital accumulation: A theory of interregional specialization and trade, *Reg. Sci. Urban Econ.,* **30**(1990), 373-404.

33. R. Tamura, Regional economies and market integration, *J. Econ. Dynam. Control,* **20**(1996), 825-845.

34. M. Weitzman, Recombinant growth, *Quart. J. Econ.,* **113**(1998), 331-360.

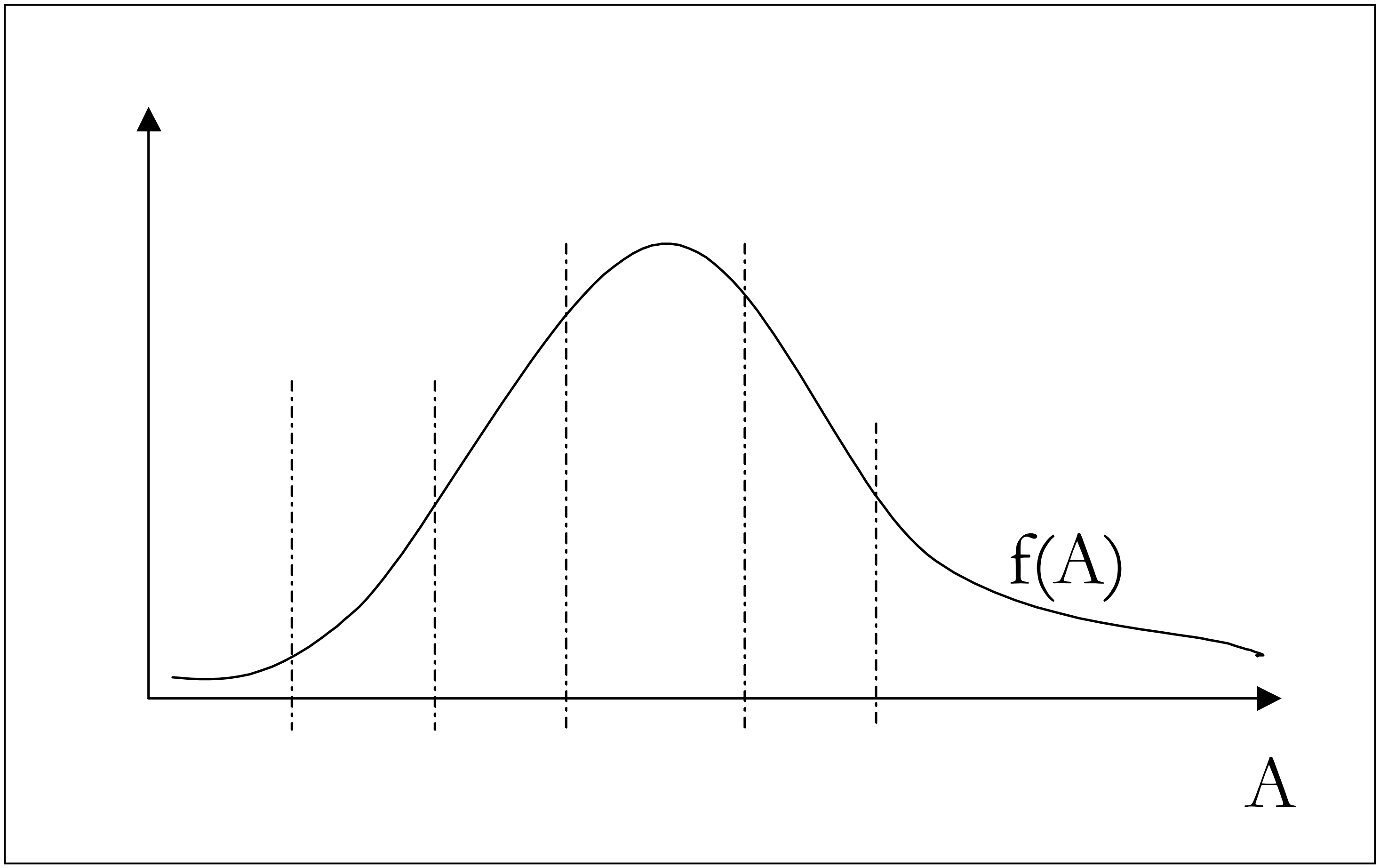


Figure 1: Clusters as Partitions Across Technology Levels

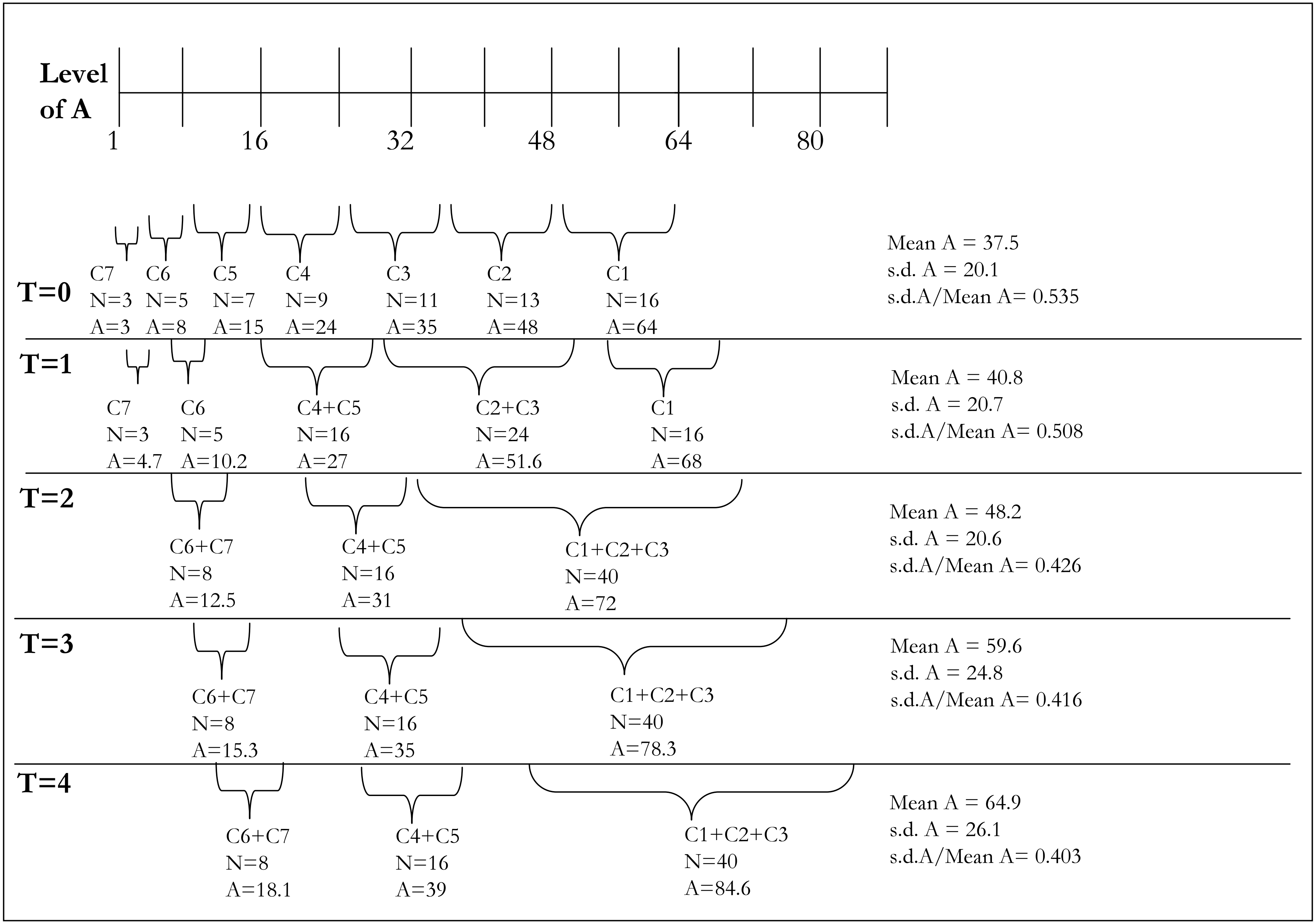


Figure 2: Example Summary

| Cluster | Size | At | Change in At | At+1 | Interaction Feasible (next up) if <1 | Technology Growth Rate | Initial | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | mean | var |
| **T=0** | | | | | | | 37.546875 | 403.8103 |
| C1 | 16.00 | 64.00 | 4.00 | 68.00 | | 0.06 | | |
| C2 | 13.00 | 48.00 | 3.61 | 51.61 | 1.06 | 0.08 | | |
| C3 | 11.00 | 35.00 | 3.32 | 38.32 | 0.99 | 0.09 | | |
| C4 | 9.00 | 24.00 | 3.00 | 27.00 | 0.99 | 0.13 | | |
| C5 | 7.00 | 15.00 | 2.65 | 17.65 | 1.00 | 0.18 | **T=0** | |
| C6 | 5.00 | 8.00 | 2.24 | 10.24 | 1.00 | 0.28 | mean | var |
| C7 | 3.00 | 3.00 | 1.73 | 4.73 | 1.02 | 0.58 | 40.82 | 429.12 |
| | | | | | | | | |
| **T=1** | | | | | | | | |
| C1 | 16.00 | 68.00 | 4.00 | 72.00 | | 0.06 | | |
| C2+3 | 24.00 | 51.61 | 4.90 | 56.50 | 0.97 | 0.09 | | |
| C4+5 | 16.00 | 27.00 | 4.00 | 31.00 | 1.95 | 0.15 | **T=1** | |
| C6 | 5.00 | 10.24 | 2.24 | 12.47 | 2.04 | 0.22 | mean | var |
| C7 | 3.00 | 4.73 | 1.73 | 6.46 | 0.99 | 0.37 | 48.22 | 422.81 |
| | | | | | | | | |
| **T=2** | | | | | | | | |
| C1+2+3 | 40.00 | 72.00 | 6.32 | 78.32 | | 0.09 | **T=2** | |
| C4+5 | 16.00 | 31.00 | 4.00 | 35.00 | 2.93 | 0.13 | mean | var |
| C6+7 | 8.00 | 12.47 | 2.83 | 15.30 | 2.00 | 0.23 | 59.62 | 615.73 |
| | | | | | | | | |
| **T=3** | | | | | | | | |
| C1+2+3 | 40.00 | 78.30 | 6.32 | 84.62 | | 0.08 | **T=3** | |
| C4+5 | 16.00 | 35.00 | 4.00 | 39.00 | 2.95 | 0.11 | mean | var |
| C6+7 | 8.00 | 15.30 | 2.83 | 18.13 | 1.99 | 0.18 | 64.91 | 684.30 |
| | | | | | | | | |
| **T=4** | | | | | | | | |
| C1+2+3 | 40.00 | 84.62 | 6.32 | 90.95 | | 0.07 | **T=4** | |
| C4+5 | 16.00 | 39.00 | 4.00 | 43.00 | 2.98 | 0.10 | mean | var |
| C6+7 | 8.00 | 18.13 | 2.83 | 20.96 | 1.98 | 0.16 | 70.21 | 757.14 |
| | | | | | | | | |
| **T=5** | | | | | | | | |
| C1+2+3 | 40.00 | 90.95 | 6.32 | 97.27 | | 0.07 | **T=5** | |
| C4+5 | 16.00 | 43.00 | 4.00 | 47.00 | 3.01 | 0.09 | mean | var |
| C6+7 | 8.00 | 20.96 | 2.83 | 23.79 | 1.98 | 0.13 | 75.52 | 833.66 |
| | | | | | | | | |
| **T=6** | | | | | | | | |
| C1+2+3 | 40.00 | 97.27 | 6.32 | 103.60 | | 0.07 | **T=6** | |
| C4+5 | 16.00 | 47.00 | 4.00 | 51.00 | 3.04 | 0.09 | mean | var |
| C6+7 | 8.00 | 23.79 | 2.83 | 26.61 | 1.98 | 0.12 | 80.83 | 913.87 |
| | | | | | | | | |
| **T=7** | | | | | | | | |
| C1+2+3 | 40.00 | 103.60 | 6.32 | 109.92 | | 0.06 | **T=7** | |
| C4+5 | 16.00 | 51.00 | 4.00 | 55.00 | 3.07 | 0.08 | mean | var |
| C6+7 | 8.00 | 26.61 | 2.83 | 29.44 | 1.99 | 0.11 | 86.13 | 997.76 |

Figure 3: Example Details

| School | D-V Adjusted pages per faculty (JEP 98) | Dept Size (Nov 00) | Public (1=yes) | Student Population Size |
|---|---|---|---|---|
| Princeton | 11.78 | 48 | 0 | 4600 |
| Harvard | 10.8 | 56 | 0 | 7087 |
| Pennsylvania | 11.03 | 39 | 0 | 9770 |
| MIT | 11.66 | 35 | 0 | 4300 |
| Northwestern | 10.37 | 49 | 0 | 7698 |
| NYU | 9.19 | 35 | 0 | 15584 |
| Boston U | 8.84 | 40 | 0 | 15470 |
| Yale | 8.59 | 49 | 0 | 5266 |
| UC SanDiego | 7.96 | 30 | 1 | 16230 |
| Stanford | 6.91 | 50 | 0 | 6594 |
| Texas-Austin | 7.05 | 30 | 1 | 36750 |
| Rochester | 8.61 | 28 | 0 | 4260 |
| UC Berkeley | 6.41 | 54 | 1 | 22678 |
| Maryland | 6.52 | 38 | 1 | 24717 |
| Pittsburgh | 6.56 | 28 | 0 | 16798 |
| Johns Hopkins | 9.14 | 15 | 0 | 3904 |
| Chicago | 7.5 | 31 | 0 | 3917 |
| Minnesota | 6.87 | 27 | 1 | 34765 |
| Wisconsin-Madison | 5.86 | 42 | 1 | 28270 |
| Virginia | 6.74 | 24 | 1 | 12467 |
| UC Los Angeles | 5.21 | 59 | 1 | 24668 |
| Columbia | 5.17 | 46 | 0 | 6000 |
| Michigan | 4.32 | 52 | 1 | 23977 |
| Ohio State | 5.03 | 45 | 1 | 42596 |
| Duke | 5.34 | 33 | 0 | 6300 |
| UC Irvine | 5.9 | 29 | 1 | 15522 |
| Brown | 4.69 | 29 | 0 | 5782 |
| Penn State | 4.1 | 26 | 1 | 40000 |
| UC Davis | 5.19 | 25 | 1 | 19460 |
| Iowa | 3.81 | 25 | 1 | 19537 |
| Florida | 4.05 | 23 | 1 | 28642 |
| Cal Tech | 5.44 | 24 | 0 | 900 |
| UC Santa Barbara | 3.86 | 34 | 1 | 16700 |
| Texas A&M | 3.62 | 30 | 1 | 36044 |
| Boston College | 3.81 | 25 | 0 | 8700 |
| Washington-Seattle | 3.14 | 29 | 1 | 25638 |
| Carnegie Mellon | 3.49 | 37 | 0 | 4600 |
| North Carolina | 2.41 | 32 | 1 | 15400 |
| Illinois-Urbana | 1.94 | 40 | 1 | 27889 |
| Vanderbilt | 2.43 | 41 | 0 | 6037 |
| Virginia Tech | 2.47 | 17 | 0 | 21415 |
| Cornell | 2.25 | 53 | 0 | 13669 |
| Colorado | 1.94 | 27 | 1 | 20595 |
| Michigan State | 1.73 | 38 | 1 | 34000 |
| Houston | 1.82 | 21 | 0 | 23525 |
| Rutgers | 1.63 | 34 | 1 | 24900 |
| N. Carolina State | 1.74 | 29 | 1 | 21990 |
| Rice | 1.86 | 21 | 0 | 2748 |
| SUNY Albany | 1.87 | 24 | 1 | 11617 |
| Georgetown | 1.58 | 27 | 0 | 6176 |

Notes:
1. Adjusted pages per faculty are from Dusansky and Vernon, 1998.
2. Department size data is from each department's website, November 16, 2000. The data does not include lecturers or adjunct faculty, but does include emeritus faculty.
3. Public refers to whether the university is a public university (1) or private (0).
4. Student population is the undergraduate population as reported by the school's website on November 17, 2000. The data are for the 1999 or 2000 academic year where specified by the website.

Figure 4: Departmental Data

| Difference in Cardinal Ranking | Difference in Department Size | Difference in Department size per student |
|---|---|---|
| 49 | 64 | 63 |
| Note: Null hypothesis is mean zero. | | |

Figure 5: Chi-Squared Tests

| Y Variable = Adjusted pages per faculty | X(1) Variable = Intercept | X(2) Variable = Department size | X(3) Variable = Public (1=yes) | X(4) Variable = Student size/Department size |
|---|---|---|---|---|
| OLS Estimate | 4.334 | 0.067 | -1.607 | -0.008 |
| Standard error | 1.717 | 0.040 | 1.023 | 0.001 |
| Overall-F Test = 4.16 | | | | |
| | | | | |

Figure 6: Regression Results