14.662 Spring 2018
Lecture Note: The Economics of Superstars *and* Mediocrities

David Autor, MIT and NBER

February 21, 2018

Introduction

We've so far considered models of skill demand and wage structure in which individuals within a skill group are *perfect substitutes*. That is, although H and L workers are imperfectly substitutable when σ is positive but finite, all H workers are perfectly substitutable with one another (and similarly for L workers). The model allows for workers to have more or less H (or L), but all endowments of H still combine linearly within the production function. This notion of perfect within-group substitutability probably fits some settings, but it certainly does not fit settings in which individuals have unique talents (e.g., athletes and performers, doctors and lawyers, scientists and technologists, academics and writers, etc.). In some sense, each individual in these markets is a 'market of one'—which would make a good title for an economics-themed Hollywood thriller—and no other actor is fully equivalent. But it would be inaccurate to imagine that each unique worker resides in a distinct market. Clearly, we choose *among* doctors according to quality, price, waiting time, etc. So it must be the case that although each is unique, we can still make tradeoffs among them. Moreover, casual empiricism suggests that the market gives rise to well behaved demand and supply relationships governing the allocation of unique talents. The general class of settings in which a group of unique factors are allocated to one another (doctors to patients, athletes to teams, CEOs to firms) is called an *assignment problem*. This lecture discusses how assignment markets work in theory.

1 Assignment Models and the Economics of Superstars

Assignment models were introduced by Nobel Laureate Jan Tinbergen in the 1950s, and they were 'popularized' (okay, made less obscure) by Sattinger's widely cited 1993 *JEL* survey. The defining feature of assignment models is the presence of indivisibilities among factors of production. In a situation where the amounts of two matching factors cannot be shifted across different units of production, factors are *not* necessarily paid their marginal products in the standard sense (more on this soon). As we will see, it's possible (indeed likely) for the wage paid to a given worker to rise or fall due to changes in the distribution of ability of other workers *without* any change in the productivity of that worker or in the value of his output.

Sattinger's 1975 *Econometrica* paper, "Comparative Advantage and the Distributions of Earnings and Abilities," is the first modern formulation of an assignment model. In this deeply insightful and original paper, Sattinger sets his sights on the following intellectual target:

> "This paper constructs a model of the allocation of workers to jobs. The intention is to find the minimum requirements for the distribution of earnings to be different from the distribution of abilities. It is not necessary to depart from the assumptions of perfect competition or marginal productivity wage determination. All that is required is that there be comparative advantage in the performance of tasks by individuals."

That paragraph captures three ideas that are foundational to this entire body of research:

1. Allocation is indivisible from the distribution of earnings in an assignment setting: what you earn *does depend* on where you work. There is not a “law of one price for skill” here. This contrasts sharply from conventional settings where workers are perfect substitutes within a skill group, so each worker’s earnings is a scalar function of her skill endowment.

2. A natural intuition from introductory microeconomics that in a competitive labor market, the distribution of wages will be directly proportional to the distribution of abilities. Sattinger is explicitly interested in discovering competitive settings where this is *not* the case.

3. The key condition for this non-equivalence between the distribution of wages and that of abilities is *comparative advantage.* In a setting with comparative advantage, earnings will depend on both ability and assignment. Assignment will magnify the importance of ability, meaning that the distribution of earnings may be far more skewed than the distribution of underlying abilities.

Assignment models have defied widespread use, perhaps in part because they are typically rather non-intuitive and analytically difficult. Sattinger’s paper is quite clearly exposited, but it’s not exactly a cakewalk to follow—intuition is left for the reader to supply. I admire Sattinger’s paper enormously, but I will not teach it in class. I will focus instead on the 2008 paper by Marko Terviö, which conveys very much the same economics with even more eloquence and intuition.

1.1 Rosen’s Superstars paper

Before turning to an assignment model, let me draw an intellectual link to the famous 1981 *AER* article by Sherwin Rosen, “The Economics of Superstars.” This paper is often cited as a prescient explanation for the rise in returns to skills that was to occur in the years immediately following its publication.[1] Rosen’s observation was that certain services face a demand structure that places considerable weight on quality versus quantity. For example, a patient would rather have one really good heart operation than two mediocre ones; an opera fan would rather see one Pavarotti concert than ten mediocre opera performances; no amount of Applebee’s food sums up to a meal at Oleana.

If there is widespread agreement on *which* providers of these services offers the highest quality, one can imagine that these providers would earn considerable rents. But there is also a natural check on the possibility that one ‘talent’ will control an entire market: congestion. A single heart surgeon cannot perform all of the surgeries; a great cook can only make so many dinners in an evening. It may be these congestion externalities that prevent single talents from controlling entire markets.

Now, imagine a technological improvement that reduces congestion, allowing holders of talent to more efficiently serve larger markets. Examples: television, recording technology, the Internet, etc. These types of technological advances may increase the rents accruing to the most talented

[1] Sherwin died from cancer in 2001 with almost no advance warning while still a youthful 61 years old. He plausibly would have won the Nobel prize in Economics had he not been lost to us so prematurely. Rosen was a great thinker as well as a stellar individual.

individuals by increasing their market share (possibly to a global scale). Some professions in which this is likely to be important: performing artists and entertainers; managers of large corporations (CEOs); athletes; fund managers; possibly some academics. Is is actually hard to make a very long list.

Rosen's insight almost surely helps to explain why athletes and movie-stars earn such enormous salaries, and why these salaries have risen as mass communications and entertainment have improved. Although I don't think that Rosen was aware of it at the time, the underlying ideas in the Superstars paper are very much in Sattinger's paper. Assignment magnifies the importance of skill in settings with comparative advantage. Rosen gave them a specific narrative—which was an invaluable contribution—and applied a different modeling approach that was probably less of a contribution.

2 Applying an Assignment Model to CEO Pay

CEO pay is a divisive topic on which there has been an enormous amount of research. There is room for debate about whether this topic receives more attention than it merits. (I personally consider it to be next door to sports economics, a topic that I hold in similarly low regard.) Irrespective of your enthusiasm for the topic, however, the 2008 *AER* paper by Marko Terviö is a brilliant contribution—not primarily for its empirical application to CEO pay (though that's quite worthy) but for its conceptual model, which studies CEO pay and productivity through the lens of an assignment model. Terviö presents a beautiful exposition of the intellectual foundation of assignment models by reformulating the basic assignment model using distributional ranks rather than actual distribution functions. This formulation is completely natural because assignment models are inherently ordinal rather than cardinal. The paper's substantive conclusions on CEO pay and its relationship to the distribution of CEO talent are also intuitive and thought-provoking (which does not mean that they are correct, though it's not evidence against that possibility!).

2.1 Basic setup

There are three assumptions that simplify the model:

1. One-dimensional quality of factors
2. Continuity of quality distributions
3. Complementarity between the qualities of factors

The first two assumptions are simply for analytical convenience. The third assumption is essential to the conceptual framework. The paper further assumes risk neutrality and symmetric information (to abstract away from agency problems and focus your attention on Terviö's main points).

Worker quality is denoted as a. Firm quality is denoted by b, which will be referred to as firm size. These are the only two factors of production in this model. There is a unit mass of workers and firms with continuous, finite support and no gaps. (The unit mass should be thought of as a normalization of all units that are active in equilibrium.)

The production function is continuous and strictly increasing in both its arguments. The key assumption is that this function has a positive cross-partial between a and b, so these inputs are complements. Accordingly, efficiency requires positive assortative matching: it would be inefficient to put a low a with a high b when they magnify one another's productivity. (Side note: If you took 14.121 and thought that the concept of *supermodularity* would never come up again, then you will be pleasantly surprised to learn that it's the key assumption driving this model.[2])

As is explained in the paper, one can write the production function *without loss of generality* as:

$$Y(a,b) = a \cdot b.$$

The notation in the paper is actually an important contribution, since it's surprisingly clean. Order abilities by quantile so that $a[i]$ is the ability of the the i^{th} quantile individual, with $a'[i] > 0$. Denoting the distribution function by F_a, the profile of a is defined by

$$a[i] = a \ \leftrightarrow \ F_a(a) = i.$$

2.2 Equilibrium conditions

There are two types of constraints that allow us to characterize an efficient allocation of workers to firms: a sorting constraint, which requires that no worker/firm pair wish to rematch; and an incentive compatibility constraint, which requires that all workers and firms earn income at least equivalent to their outside option:

$$\begin{aligned} Y(a[i], b[i]) - w[i] &\geq Y(a[j], b[i]) - w[j] \quad \forall i,j \in [0,1] \qquad SC(i,j) \\ Y(a[i], b[i]) - w[i] &\geq \pi^0 \quad \forall i \in [0,1] \quad PC\ b[i] \\ w[i] &> w^0 \quad \forall i \in [0,1] \quad PC\ a[i] \end{aligned}$$

It is assumed that π^0 and w^0 are the same for all units, though it would be sufficient to assume that outside opportunities increase slower along the profile than do equilibrium incomes. The lowest active pair breaks even:

$$Y(a[0], b[0]) = \pi^0 + w^0.$$

If there were n workers and n firms, there are in theory $2n!$ sorting constraints, where n is the number of units. However, most constraints are redundant since for $i \geq j \geq k$, $SC(i,j) + SC(j,k)$ implies $SC(i,k)$. The binding constraints are therefore: (1) the marginal sorting constraints that keep firms from wanting to hire the next best individual; and (2) the participation constraints of the lowest types.

[2] For the geeks in the room, an noteworthy difference between Terviö's model and Sattinger's model is that Sattinger's setup invokes log supermodularity. Sattinger didn't invoke that term, which had not been invented at that time—or at least hadn't been adopted by economists. But that makes his paper all the more 'super' IMHO.

Ordering the binding sorting constraints and returning to our continuum of i (so we can use calculus), one can write:

$$\frac{Y\left(a\left[i\right],b\left[i\right]\right)-Y\left(a\left[i-\varepsilon\right],b\left[i\right]\right)}{\varepsilon}\geq\frac{w\left[i\right]-w\left[i-\varepsilon\right]}{\varepsilon}. \tag{1}$$

This equation becomes an equality as $\varepsilon\rightarrow 0$, and yields the slope of the wage profile:

$$w'\left[i\right]=Y_a\left(a\left[i\right],b\left[i\right]\right)a'\left[i\right],$$

where Y_a is the partial derivative. To get the full wage profile, we need to then integrate over the production function and add in the binding participation constraint:

$$w\left[i\right]=w^0+\int_0^i Y_a\left(a\left[j\right],b\left[j\right]\right)a'\left[j\right]d\left[j\right].$$

Note that this set of sorting conditions could equivalently be written in terms of workers choosing firms rather than firms choosing workers; these cases are isomorphic. So, similarly, the profile of profits must satisfy:

$$\begin{aligned}\pi'\left[i\right] &= Y_b\left(a\left[i\right],b\left[i\right]\right)b'\left[i\right]\\ \pi\left[i\right] &= \pi^0+\int_0^i Y_b\left(a\left[j\right],b\left[j\right]\right)b'\left[j\right]d\left[j\right].\end{aligned}$$

These conditions also imply that $y=\pi+w$ at each firm.

This is an extremely tight set of constraints on the problem. It implies that the wages and profits of each factor depend at quantile i depend on the full profile of factors from quantile 0 to $i-\varepsilon$ (but not on the profile above i).

A number of useful observations follow (and you should read the paper, which is chock full of astute insights, and even more so in the footnotes):

1. There is no bargaining in this model because of the continuity of the distribution of both factors. If there was a jump at some point in the profile of one factor, then all of the surplus would go to the factor with the jump because the other side is perfectly competitive.

2. Payments to factors are only affected by the quality of those *below* them in the ranking. This is because the binding constraint on each worker or firm is that the quality/price of the worker/firm just below it in the distribution.

3. The unobserved productivity characteristics a and b are essentially ordinal. Any increasing transformation of the scale of measurement for a factor's quality combined with the inverse change in the functional form of the production function changes nothing of substance in the model. In this sense $Y\left(a\left[i\right],b\left[i\right]\right)=a\left[i\right]\times b\left[i\right]$ is a general functional form (so long as we are assuming complementarity).

4. Quoting Tervio, "It would be incorrect to say that factors earn their marginal productivity by the usual definition of marginal productivity, because the increase in output if the individual of ability $a[i]$ were to increase in ability is proportional to $b[i]$. But if she were to increase in ability, then, in equilibrium, she would also move up in the ranking and be matched with a higher b— and other individuals would have to move down and experience a decrease in productivity. This means that individuals in fact do receive their marginal product, once the margin is defined correctly. As ability cannot conceivably be extracted from one individual and poured into another, the relevant margin here is whether an individual will participate in the industry or not—and if not, then the effect of the resulting rearrangement of remaining individuals is part of the marginal product."

2.3 Comparative statics

Consider a change by which the production function Y is multiplied by some constant G with no change in distributions of a and b. Uniform productivity growth is mathematically equivalent to changing the units of measurement for output. This gives rise to a useful lemma:

Scaling lemma: If $Y_t(a,b) = GY(a,b)$, $w_t^0 = Gw^0$ and $\pi_t^0 = G\pi$, then $w_t[i] = Gw_t[i]$ and $\pi_t[i] = G\pi_t[i]$ for all $i \in [0,1]$.

Figure 1 summarizes much of the intuition of the model. In the strictly multiplicative case, the output accruing from matching a worker of ability a and a firm of ability b is $a \times b$, which is simply a rectangle in a Cartesian graph. This graph is drawn as follows. Let $a = \varphi(b)$, defined by $a[F_b(b) = \{(a,b)\ \ st.F_a(a) = F_b(b)\}]$ with slope

$$\varphi'(b) = a'[F_b(b)]\, f_b(b) = \left.\frac{a'[i]}{b'[i]}\right|_{i=F_b(b)}.$$

$\varphi(b)$ is strictly increasing in b, and the slope is given by the relative steepness of a and b at each quantile i.

Be sure to understand Figure 1, since essentially all of the intuition of the assignment model is contained in this figure.

Here's an interesting exercise to build intuition. Let's check Terviö's assertion that "the relevant margin here is whether an individual will participate in the industry or not—and if not, then the effect of the resulting rearrangement of remaining individuals is part of the marginal product." Consider a hypothetical case where the highest ability worker (index $a[1]$) falls in ability to that of the lowest ability worker $a[0]$. Can we demonstrate that the reduction in total output is equal to this worker's wage?

- The "demotion" of the highest ranked worker means that each firm other than $b[0]$ will have to match with a slightly lower ranked worker. So, previously, total output was equal to:

$$Y = \int_0^1 Y(a[j], b[j])\, d[j]$$

- Now, each firm except for the lowest ranked firm $b[0]$ will have to pair with a slightly lower quality worker. Thus, the fall in output at each firm $Y[j]$ is

$$\Delta Y[j] \equiv Y[j] - \hat{Y}[j] = Y(a[j], b[j]) - Y(a[j-\varepsilon], b[j]). \tag{2}$$

The output of the lowest ranked b is unaffected since it still pairs with the lowest ranked a. But for all units b above rank 0, they will now be pairing with a units that are ε below them in the prior ranking. Note that $a[\cdot]$ continues to refer to the values of a in the *original* distribution, not the new distribution.

- Dividing equation (2) by ε and letting $\varepsilon \to 0$, we take the limit of

$$\frac{\Delta Y[j]}{\varepsilon} = \frac{Y(a[j], b[j]) - Y(a[j-\varepsilon], b[j])}{\varepsilon},$$

to get

$$Y'[j] = Y_a(a[j], b[j])\, a'[j].$$

- We integrate over the full distribution of units to obtain the total loss in output:

$$\Delta Y = \int_0^1 Y_a(a[j], b[j])\, a'[j]\, d[j]$$

This is the net reduction in output caused by worker $a[1]'s$ demotion.

- Having solved for the reduction in output, let's compare this to the change in the wage bill. The wage of the previously highest ability worker falls from

$$w[1] = w^0 + \int_0^1 Y_a(a[j], b[j])\, a'[j]\, d[j]$$

to

$$\hat{w}[1] = w^0.$$

Thus the change in the wage for worker $a[1]$ is

$$\begin{aligned} w[1] - \hat{w}[1] &= w^0 + \int_0^1 Y_a(a[j], b[j])\, a'[j]\, d[j] - w_0 \\ &= \int_0^1 Y_a(a[j], b[j])\, a'[j]\, d[j], \end{aligned}$$

which is identical to the fall in total output, ΔY. This confirms Terviö's assertion.

2.4 Applying this model to CEOs

There are numerous conceptual and practical issues to confront before bringing this model to the data. I will discuss these assumptions and results only briefly.

The issues:

1. The surplus created by the CEO-firm interaction is unobserved. The market value of the firm is affected by the current CEO and by expectations of future productivity (which also depend on future CEOs)—thus firm size, an outcome, cannot be treated as the b variable.

2. A second issue is that part of the market value of the firm will surely reflect the value of capital that can be readily transferred/resold among firms. This capital is *not* indivisible and so is not part of the surplus represented in $Y(\cdot)$.

3. A third issue is that the current market value of a firm depends on the quality of the current CEO and the quality of past and (in expectation) future CEOs. Thus, one cannot necessarily infer the current CEO's quality purely from contemporaneous data.

4. Productivity tends to grow over time, and the expectation of growth further affects current market value.

5. The distribution of CEO ability and latent (exogenously determined) firm size (not market value) can change over time.

6. Outside options may shift.

Thus, the subsequent analysis rests on a number of strong assumptions that necessarily make the exercise speculative (though perhaps still informative if one believes the foundational assumptions of the assignment model):

1. The distribution of $a[i]$ and $b[i]$ are time invariant.

2. Productivity grows deterministically at rate g at all firms (that's why the scaling lemma is needed)

3. The value of outside options grow at rate g. Tervio uses values between 0.2 and 0.025.

4. The discount rate is constant. Tervio uses values between 0.08 and 0.05.

5. The impact of past *and future* CEO quality on current firm performance decays at a constant rate $\alpha_{\tau+1} = \alpha_\tau \lambda / (1+\lambda)$. λ determines the decay rate. With $\lambda \to \infty$, only the current CEO affects contemporaneous earnings. Tervio uses values between ∞ and 0.1

6. Since adjustable capital must earn the market rate of return, it must be subtracted from Y. To determine the contribution of adjustable capital, Tervio assume's that the gross surplus has constant elasticity θ with respect to adjustable capital. Tervio sets the share of adjustable capital in Y at values between 0 and 0.8.

These assumptions are sufficient to pin down the contribution of firm size and ability to output up to an additive constant (one constant each for the two factors). The reason the constant is indeterminate is that one cannot infer the contribution of firm size and ability to surplus at the smallest firm in the sample since this depends on the full unobserved distributions of size and ability below this firm. However, these constants drop out when estimating counterfactuals of the form 'how much would surplus rise or fall *relative to* the current baseline if the distribution of CEO ability were modified in various ways?'

Terviö explores three counterfactuals for CEO ability: all firms are paired with a CEO of the ability of the lowest ranked CEO; all firms are paired with the median CEO; all firms are paired with the highest ranked CEO. One can also perform a similar counterfactual for firm size. See Tables 1 and 2. Several points:

1. The contribution of firm size to CEO rents is much larger than the contribution of CEO ability. That is, in the logic of this model, the high levels of CEO pay at the top are mostly due to the exogenous component of firm scale. A little extra skill can go a long way when you have a lot of people (or capital) working for you.

2. The distribution of CEO ability *relative to* the lowest ability CEO can be inferred from the differential equations that describe the equilibrium outcomes. See Figure 5. A key conclusion, and one that should be intuitive, is that the distribution of CEO ability is *not* wide. It is estimated that the CEO at the top ranked firm is only 5 to 8 percent more able than the CEO at the bottom ranked firm.

3. To see the role of firm size in CEO pay, consider figure 4 where all firms are replaced with the lowest, median or highest ranked firm while the CEO ability distribution is held constant. Panel B is perhaps most informative. If all firms were the median size, the top CEO would earn 12.5 million less than the 500th CEO than is currently the case and the bottom CEO would earn an additional 2.5 million less relative to the 500th CEO than is currently the case. (Why greater pay dispersion at the bottom? Because the flat profile of firm size, all surplus is allocated to workers–unlike in the continuous firm distribution case. This surplus redistribution effect can offset the leveling effect of greater equality of firm size. The relevant comparison here is not really among CEOs in these various scenarios but rather between scenarios–highlighting how changes in firm size hugely affect the level and dispersion of CEO pay, holding CEO talent constant.)

2.5 How the heck does Terviö calculate what part of the capitalized market value of firms is due to CEO vs. firm 'quality?'

This is pretty tricky.

- Under the assumptions of (1) deterministic, constant growth; (2) proportionate growth in the value of outside options; and (3) no change in the quality distribution of factors over time, one obtains the following relationship

$$\frac{w[i]}{1-B} + \nu[i] = \frac{a[i]\,b[i]}{1-B},$$

where $B = (1+g)/(1+r)$ is the growth adjusted interest rate. The first term on the left is the present value of pay to *all* CEOs that firm i will ever employ and $\nu[i]$ is the market value of the firm. As shorthand, Terviö refers to $1/(1-B)$ as the price to earnings ratio.

- Adjustable capital can be partialed out from $\nu(i)$ using a Cobb-Douglas assumption. Only the part of ν that excludes adjustable capital is subject to the assignment process.

- The differential equations (24) and (25) in the paper permit inference on the quality of CEOs and firms at each percentile in the distribution *relative to* the lowest quality units $(a[0], b[0])$ participating in the market:

$$\frac{a[i]}{a[0]} = \exp\left\{\frac{\lambda}{\lambda+1-B}\int_0^i \frac{w'[j]}{w[j]+v[j](1-B)}dj\right\},$$

$$\frac{b[i]}{b[0]} = \exp\left\{\int_0^i \frac{v'[j]-w'[j]/\lambda}{w[j]+v[j](1-B)}dj\right\},$$

  where λ is the decay rate of present CEO contributions to future firm value.

- I think a large part of the intuition of Figure 5 is that CEO pay increases very, very little from the smallest to largest firm, whereas market value increases by hundreds of billions (say from 19 billion at rank 1,000 to 750 billion at rank 1,000). This loosely implies that the slope of CEO talent must be relatively shallow.

2.6 Conclusions

You do not have to be convinced that this model or empirical exercise fully (or even partly) explains the level or trends of CEO pay to find much of value in this paper. As a labor economist, you should add the basic workings of assignment models to your intellectual toolkit. And this is best paper I've seen for acquiring those tools. As I will discuss in class, the assignment model can be used to interpret the Rosen 1981 *AER* superstar's model, though Rosen's original model is not constructed in the form of an assignment model.

I will not spend class time on the related paper by Gabaix and Landier (*QJE* 2008) and the critique of Gabaix-Landier by Gordon and Dew-Becker ("Unresolved Issues in the Rise of American Inequality," *Brookings*). To their credit, Gabaix and Landier also recognized how the assignment model could be used to explain CEO pay and arguably attempted a more ambitious calibration than did Terviö. People who study CEO pay over longer time intervals and across nations have generally not been persuaded by their evidence (see for example Frydman and Saks, 2010[3]).

[3] Frydman, Carola and Raven E. Saks. 2010. Executive Compensation: A New View from a Long-Term Perspective, 1936-2005. *Review of Financial Studies*. 23: 2099-2138.

3 Terviö: Mediocrity in Talent Markets

Having nailed the 'superstars' model in his *AER* paper, Terviö flips it on its head in his 2009 *ReStud* paper, "Superstars and Mediocrities: Market Failure in the Discovery of Talent." This highly original paper casts the superstars phenomenon in a fresh but less favorable light. The point of departure between the Sattinger/Rosen assignment models and the Terviö mediocrity model is in the mechanism in which talent is discovered. Sattinger/Rosen take it as given that talent is *known* to the market. Given full information, the wages paid to talent are likely to be efficient.

Terviö's 2009 paper takes a step back by asking: how does talent become known? Terviö observes that to be 'discovered,' one must first 'audition.' This observation seems plausible for many talent-occupations. In the Terviö model, knowledge about worker quality is a joint output of production. (e.g., to find out if a worker would make a good actor, he needs to appear in a movie.) Economist Richard Caves refers to this ex ante uncertainty about talent as the "Nobody Knows" property—one cannot evaluate talent without putting it to use. This implies that there is real resource cost to discovering talent. This resource could be capital or, more realistically, it could be the time of skilled evaluators (talent agents, moviegoers, academics, passengers on an airplane, etc.).

Now add a second broad observation: the person (or firm) that discovers the talent does not necessarily have the ability to capture the full value of that talent discovery. The facts that (1) there are real costs to discovering talent but (2) the rewards to those discoveries may not be entirely appropriable, suggest that this setting may generate market inefficiencies.

3.1 Setup

- Firms need one worker and one unit of capital at cost $c > 0$ to produce output in each period.
- Output of a firm is equal to the worker's talent: $Y = \theta$.
- Talent is drawn from a distribution with a continuous and strictly increasing cumulative distribution function, F, with positive support $[\theta_{min}, \theta_{max}]$.
- There is an unlimited supply of potential workers, all of who have an outside wage of w_0. Hence, *talent is not scarce.*
- Talent is industry-specific and becomes public knowledge after one period of work; the worker may then work in the industry up to T more periods, after which he ceases to be productive.
- Both workers and firms are risk neutral, and there is no discounting.
- Industry output faces a downward-sloping demand curve $p^d(q)$. We think of the "quantity" of talent as referring to the quality of output times the number of workers I. Thus, all else equal, the equilibrium price of output is lower when talent in the industry is higher (i.e., less scarce).

- The number of firms is "large" so that individual firms have no impact on total output, and there is no uncertainty about the realization of the distribution of talent.
- Let the number of firms (and jobs) be a continuous variable I equal to the mass of the industry workforce, normalize so that $I = 1$
- Long-term wage contracts are not enforceable; workers cannot commit to decline higher offers from other firms in the future.

Some key mechanics

- The planner (or the market) will determine a talent threshold ψ above which workers will be retained in the industry and below which they will not.
- The social planner would like to maximize average talent in the industry. This is because the opportunity cost of talent in the outside market is uncorrelated with talent. Thus, the higher the level of talent *in* the industry, the higher is aggregate social welfare. (Note that output is linear in talent, so the average is a sufficient statistic for the maximization problem.)
- In steady state, inflows of talent (new workers, that is) must equal outflows of talent.

Some key economics

In an efficient market allocation

1. The veteran worker of the lowest type ψ retained in the industry must be indifferent between exiting and remaining in the industry, and thus must earn exactly the outside wage w_0
2. Novices must be indifferent between entering the industry or pursuing the outside career, taking into account their option to exit for the outside career later on (that is, before $T + 1$ periods have elapsed)
3. Firms must expect zero profits from hiring any talent, since they are assumed to be competitive
4. The output price must clear the market

3.2 Characterizing the talent threshold

- In both market equilibrium and in the social planner's optimal solution, individual careers will proceed in a simple manner:
  - *After one period of work, those whose talent is revealed to be below a certain threshold level ψ will exit the industry, while those above the threshold ψ will stay on for T more periods. This exit threshold will be the key variable in the model.*

- In steady state, **inflows must equal outflows**. Denote the fraction of novices in the workforce by i; a fraction $F(\psi)$ of them exit. The remaining fraction of jobs $1-i$ are held by veterans; a fraction $1/T$ of these, the oldest cohort, retires each period. Equating the flows of exits and entries yields:

$$iF(\psi) + \frac{1}{T}(1-i) = i$$
$$i(\psi) = \frac{1}{1+T(1-F(\psi))}$$

- **Average talent in the industry is therefore**

$$A(\psi) = i(\psi)\bar{\theta} + (1-i(\psi))E[\theta|\theta \geq \psi].$$

One can substitute the equation above for $i(\psi)$ into this expression to obtain

$$A(\psi) = \frac{1}{1+T(1-F(\psi))}\bar{\theta} + \frac{T(1-F(\psi))}{1+T(1-F(\psi))}E[\theta|\theta \geq \psi].$$

Notice that a higher ψ means that a larger fraction of workers are novices (which lowers average quality) but the average quality of retained veterans is higher, which raises average quality.

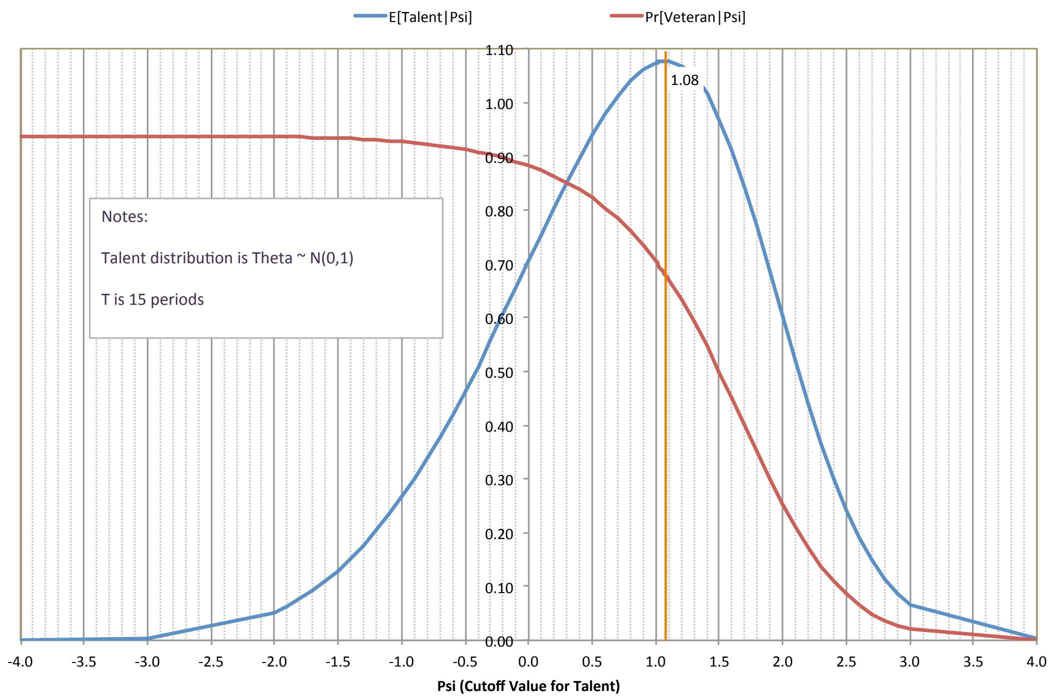
- The planner's choice of ψ: *This expression is maximized when the average talent in the industry (averaging over new entrants and veterans) is equal to the marginal talent retained.* (When ψ is too low, raising it increases both the marginal and the average; when ψ is too high, raising it raises the marginal but reduces the average because there are too few veterans.)

- The market equilibrium condition determining ψ can be written as

$$\psi - \bar{\theta} = T(1-F(\psi))(E[\theta|\theta \geq \psi] - \psi), \tag{3}$$

the solution to which is denoted as A^*. The LHS of this expression is the foregone output of hiring a novice rather than a veteran (invoking that average output is equal to marginal output ψ at A^*). The RHS is the expected future gain from hiring a novice, equal to the probability that the novice is retained for T additional periods times expected output (conditional on being retained) in each subsequent period (invoking that ψ is constant across future periods).

- The paper demonstrates that the solution to this market equilibrium condition is unique with no distributional assumptions on θ required (except continuity and boundedness). Here's a numerical example of the optimal choice of ψ with $\theta \sim N(0,1)$ and $T = 15$:



3.3 Indifference condition for entrants and veterans

- Veterans must choose to exit if $w(\theta) < w_0$, therefore the marginal veteran is indifferent between exiting and remaining in the industry

$$w(\psi) = w_0.$$

- Similarly, novices must be indifferent about entering the industry, accounting for the option to exit if $\theta < \psi$:

$$w\left(\bar{\theta}\right) = T \times E\left[\max\left\{w(\theta), w_0\right\}\right] = (1+T)\, w_o.$$

  This expression may imply that $w\left(\bar{\theta}\right) < 0$, that is entrants will receive a *negative* wage. The optimal solution may therefore require an absence of binding credit constraints.

- If constrained individuals cannot accept a wage below $w_0 - b$ where $b > 0$, then the wage of novices will be given by

$$w\left(\bar{\theta}\right) = w_0 - b.$$

  This will yield substantial market-level inefficiencies as we shall see.

3.4 Zero profit condition

- Firms must expect zero profit from hiring any talent. Denoting the output price as P, this implies that
$$P\theta - c - w(\theta) = 0.$$

- Since there is no risk aversion and information is symmetric, novices can be treated as a known quantity with talent $\bar{\theta}$.

3.5 Market clearing

- The output price must clear the market. We've stipulated downward sloping demand, and we know that average talent is rising in ψ up to the optimum A^*. Therefore, there will be a fixed point for the market price equation where
$$P = p^d (I \times A(\psi)),$$
where $I \times A(\psi)$ is total talent supplied.

- Since talent in the industry is a function of ψ, we can rewrite the wage function in terms of individual talent and the market talent cutoff $w(\cdot) = w(\theta|\psi)$.

- We can use the zero profit condition to get an expression for P in terms of the wage.
  - A firm employing the threshold obtains revenue $P\psi$ and pays costs $w(\psi|\psi) = w_0 + c$.
  - Thus, the equilibrium output price must satisfy
$$P(\psi) = \frac{w_0 + c}{\psi}. \tag{4}$$

- Since firms must be indifferent about hiring any level of talent, the surplus must go to the worker, implying that $w(\theta|\psi) = w_0 + \{P\theta - (w_0 + c)\}$. Rearranging:
$$w(\theta|\psi) = w_0 + \left[\frac{\theta}{\psi} - 1\right](w_0 + c). \tag{5}$$
Thus, a worker at threshold talent of ψ earns w_0, which is her opportunity cost. A worker with $\theta > \psi$ earns a wage in excess of w_0 while a worker with $\theta < \psi$ would exit the market and take w_0 instead.

3.6 Solution with unconstrained workers

- Competitive equilibrium with unconstrained individuals is socially efficient, so the social planner's solution already tells us that the exit threshold must be $\psi = A^* - \bar{\theta}$.

- Inspection of the wage equation (5) reveals that novices must accept less than the outside wage w_0, since they have a positive probability of earning talent rents in the future while in the worst case they get the outside wage. If the novice wage were not below w_0, novices would earn rents and more novices would enter, contradicting equilibrium. (Concretely, the fact that the novice wage is below ω_0 is evident from the fact that $\left(\bar{\theta}/\psi\right) < 1$.)

- Market equilibrium pins down the wage function as $w\left(\theta|A^*\right)$ and $P^* = \left(w_0 + c\right)/A^*$. Intuitively, note that unconstrained individuals bid for the chance to enter the industry up to the expected value of lifetime talent rents. Since veterans of the threshold type are available at the outside wage, novices have to pay $P \times \left(\bar{\theta} - \psi\right)$ for this first period job, which exactly compensates a novice-hiring firm for its expected revenue loss (compared to what it would get by hiring a threshold type).

- In equilibrium, this novice payment must equal the expected lifetime rents: with threshold ψ , a novice has a probability $1 - F\left(\psi\right)$ of being retained, in which case he gets the excess revenue $P \times \left(\theta - \psi\right)$ as a rent on each of the T remaining periods of his career. So, the market equilibrium condition is

$$P \times \left(\theta - \psi\right) = P \times T\left(1 - F\left(\psi\right)\right)\left(E\left[\theta|\theta \geq \psi\right] - \psi\right),$$

  where the P cancels out and hence this is identical to (3) above.

- Payments by unconstrained novices raise the exit threshold to the efficient level. Thus, the inability of workers to commit to long-term contracts does *not* cause problems *if* they are able to buy their jobs upfront.

- The unconstrained payment (the price of a job) reflects the economic cost of the efficient level of experimentation.

$$b^* = P^*\left(A^* - \bar{\theta}\right) = \left(w_0 + c\right)\left(1 - \frac{\bar{\theta}}{A^*}\right)$$

  As Terviö notes, the fraction of the total costs of production, $w_0 + c$, that should be financed by the novice is increasing in $A^*/\bar{\theta}$, which is a measure of the upside potential of novices. (For small values of b^* the novice payment would merely be a wage discount below the outside wage.)

3.7 Equilibrium with constrained workers

- Suppose now that the ability of individuals to pay for their first period job is constrained at some $b < b^*$ due to an exogenous liquidity constraint. Now the novice payment is $w\left(\theta|\psi'\right) = w_0 - b$, where ψ' denotes the threshold in the constrained case.

- Then

$$\psi'(b) = \begin{cases} \left(\frac{w_0+c}{w_0+c-b}\right)\bar{\theta} & \text{for } b < b^* \\ A^* & \text{for } b \geq b^* \end{cases} \tag{6}$$

- The exit threshold is increasing in b. It follows that the average talent in the industry is also increasing in b. When novices cannot "subsidize" their employers, then the price of output must adjust upwards to induce the hiring of novices into the industry.

- This implies that when there are liquidity constraints faced by novices:

  1. threshold quality will be too low
  2. turnover will be too low and average career lengths will be too long
  3. average quality will be too low
  4. output price will be too high
  5. those who enter the industry earn rents in expectation
  6. top wages will be 'too high' because the top talents will still be drawn from the same distribution as in the unconstrained case, but they'll face a higher output price.

- I won't go through the specifics of the solution in this setting, but they are straightforward provided that you are clear on the above. Key point: *Credit constraints yield, in equilibrium, mediocre veterans in the market*

$$\psi' \in \left(\bar{\theta}, A^*\right)$$

4 Two examples

Let's work with two simple examples. Both will have the same parameter values, but in one case we'll assume that workers are not credit constrained and in the other case, we'll assume that they are. For concreteness, let's say that this industry is the movie business. The parameter c refers to the cost of making a movie. Everything is measured in $1,000's$ of dollars. So let's say that $c = 4,000$, that is, the cost of making a movie is $4 million.

Parameter values

- $T = 15$ (where the novice period does not count in T)
- Fixed cost of production is $c = 4,000$
- Workers' outside option is $w_0 = 0$
- The talent distribution is uniform on 0 to 100. $\theta \sim U[0,100]$, which implies that $\bar{\theta} = 50$ and $E[\theta|\theta \geq \psi] = 50 + \psi/2$

- Output is equal to talent

$$Y(\theta) = \theta$$

- Surprisingly, this is all the info needed to solve the model.

4.1 No credit constraints case

Start with the case where there are no credit constraints, meaning that the first period wage in the industry can be negative. Solving for equilibrium:

- Let's start with the talent threshold (equation 3)

$$\psi - 50 = 16\left(1 - \frac{\psi}{100}\right)\left(50 - \frac{\psi}{2}\right).$$

  This generates a quadratic equation with two roots, one of which is $\psi = 80$ and the other of which is out of bounds (430/3). Hence, $\psi = 80$.

- We can now obtain the wage function

$$w(\theta|\psi) = w_0 + \left[\frac{\theta}{\psi} - 1\right](w_0 + c)$$
$$w(\theta|80) = \left[\frac{\theta}{80} - 1\right]4,000$$
$$w(\theta|80) = (\theta - 80)\,50$$

  This function implies that $P = 50$. We know this because the worker obtains all of the surplus in excess of the threshold. So if the wage is increasing by 50 per unit θ, then output must be priced at 50.

- Another way to obtain P is to invoke the zero profit for hiring a marginal incumbent as in equation (4):

$$P(\psi = 80) = \frac{w_0 + c}{\psi} = \frac{0 + 4000}{80} = 50$$

- Now we need to solve for b, the cost that a novice pays to enter. From above, and using the fact that the marginal worker quality equals average worker quality, so $A^* = \psi = 80$.

$$b^* = P^*\left(A^* - \bar{\theta}\right) = (w_0 - c)\left(1 - \frac{\bar{\theta}}{A^*}\right)$$
$$b^* = 50\,(80 - 50) = -4000\left(1 - \frac{50}{80}\right) = 1,500$$

  Novices pay $1.5 million for the their first job. Why is the worker paying to enter? Because if it turns out that $\theta_i > \psi$, the worker will earn talent rents in the subsequent T periods. Those rents *must* be zero in expectation, or else entry is inefficient: novices are not indifferent about entering, meaning they are paid too much, so other distortions must result.

- *What is the probability of a novice becoming a veteran?* It's $\Pr[\theta \geq \psi] = 0.20$
- *What fraction of active workers are novices?* This is

$$i(\psi) = \frac{1}{1 + T(1 - F(\psi))} = \frac{1}{1 + 15(1 - 0.8)} = 0.25$$

- *What are the expected career earnings of a worker who exceeds the talent threshold?*

$$15 \times (50 \times \{E[\theta|\theta \geq 80] - 80\}) = 7,500$$

- Expected lifetime earnings (rents) for a novice are therefore

$$\Pr[\theta \geq \psi] \times 7,500 - b = 0.20 \times 7,500 - 1,500 = 0.$$

- And finally, the top wage in the industry is

$$w(100|\psi = 80) = 50 \times (100 - \psi) = 1,000.$$

Thus, the top earner receives **one million dollars** per movie (yes, Austin Powers joke).

4.2 Equilibrium with constrained workers

- Now assume that $b = 0$. Workers cannot pay to enter the industry. Optimality no longer applies.
- One can calculate ψ' by using (6)

$$\psi'(b) = \begin{cases} \left(\frac{w_0+c}{w_0+c-b}\right)\bar{\theta} & \text{for } b < b^* \\ A^* & \text{for } b \geq b^* \end{cases}$$

With $b = 0$, $\psi'(0) = 50$.

- Why is the talent threshold equal to mean talent in this example? Because novices cannot bid for jobs, veterans with $\theta \geq 50$ will always generate (in expectation) weakly more revenue for firms than will novices. This means that vets with $\theta \geq 50$ will always out-compete novices for a job. Thus, the talent threshold falls to the mean of novices.

- We can solve the wage equation as above

$$w(\theta|\psi) = w_0 + \left[\frac{\theta}{\psi'} - 1\right](w_0 + c)$$
$$w(\theta|\psi) = \left[\frac{\theta}{50} - 1\right] \times 4,000$$
$$= [\theta - 50] \times 80$$

This implies immediately that $P' = 80$

- We can also invoke the zero profit condition to obtain P (though the equation above already implies that $P = 80$):

$$P' \times \bar{\theta} = 4,000.$$

Since $\bar{\theta} = 50$, this implies that $P' = 80$.

- What is the expected probability of being retained after the first period? It's $\Pr[\theta \geq \psi'] = 0.50$

- The average output of veterans is $E[\theta|\theta \geq 50] = 75$, since veterans are retained if their talent exceeds the mean.

- The fraction of novices in the industry is

$$i'(\psi) = \frac{1}{1 + T(1 - F(\psi'))} = \frac{1}{1 + 15(1 - 0.5)} = 0.118$$

- The expected career earnings of a worker who exceeds the talent threshold is:

$$15 \times (80 \times \{E[\theta|\theta \geq 50] - 50\}) = 30,000$$

- Expected lifetime earnings (rents) for a novice are therefore

$$\left(1 - \frac{50}{100}\right) \times 30,000 = 15,000.$$

- And finally, the top wage in the industry is

$$w\left(100|\psi' = 50\right) = 80 \times \left(100 - \psi'\right) = 4,000.$$

Thus, the top earner receives **four million dollars** per movie (no Austin Powers joke).

4.3 Comparison

Let's compare the characteristics of these two equilibria

	Constrained Equilibrium	Efficient Benchmark
Talent Threshold ψ	50	80
Proportion Novices i	12%	25%
$E\,[\text{Talent}\vert\,\text{Veteran}] = E\,[\theta\vert\theta \geq \psi]$	75	90
Average talent A^*	72	80
Output price P	\$80K	\$50K
Top Wage $P \times (100 - \psi)$	\$4 million	\$1 million
E[Career Earnings| Veteran]	\$30 million	\$7.5 million
Wage of Novice	\$0	-\$1.5 million
E[Rents| Novice]	\$15 million	\$0

What is quite striking here is that the constrained equilibrium has *lower talent and higher wages* than the efficient benchmark. The reason is that the constrained equilibrium has artificial talent scarcity. Because talent rents accrue to workers and not firms, firms have insufficient incentive to create new opportunities for talent revelation (in fact, they have none). So, there is under-experimentation, excess retention of mediocre talent, and inefficiently high rents.

This problem could be solved in either of two ways: (1) create binding indentured servitude contracts such that a worker's wage is set independent of output and the worker cannot quit once hired (but can be fired in the 1st period) – this used to happen in European Soccer through the transfer fees system; (2) eliminate credit constraints so that workers can bid for jobs. This bidding will generate an efficient market in talent revelation. You might ask yourself whether the markets for CEOs, athletes, movie stars and other celebrity professionals appears to satisfy (1) or (2).

5 Conclusions

The key insight of the mediocrities model is that talent can receive large rents even if that talent mediocre. The reason is that the (ex-post) realization of *known* talent may have artificial scarcity. An interesting point on which to speculate is whether this model is, in some sense, more general than the Rosen model; could one could potentially nest the Rosen model as a special case where talent revelation was costless? Almost surely the answer is yes. It will be interesting to see whether this paper move's the economic prior that the market for superstars is in some sense socially efficient.

Talent, comparative advantage, talent assignment, and talent *discovery* are fascinating areas of economic inquiry that, in academic turf terms, are populated by a small number of relatively brilliant papers. The most recent such entry is Amanda Pallais' 2014 *AER* paper, which takes as its inspiration Terviö's 2009 *ReStud* paper. As you're likely aware, Mandy is a 14.662 alumna, and the intellectual seed for her job market paper (AKA the 2014 *AER* paper) took root in this class.