Increasing Returns to Scale Within Limits: A Model of ICT and its Effect on the Income Distribution and Occupation Choice

Tianxi Wang and Greg C. Wright*

August 2020

Abstract

A key feature of Information and Communication Technologies (ICT) is that they increase the size of the market – or the "scale of operation" – for workers in some occupations. We model the scale of operation as the limit up to which the production technology displays increasing returns to scale. We then explore the implications of this feature of ICT for the income distribution within affected occupations, as well as for individuals' occupational choices. Within occupations, an increase in the scale of operation intensifies competition between workers and increases inequality. It also drives the lowest-ability workers out of the occupation while reducing the earnings of the next lowest-ability workers when the substitutability between the output of the affected occupations and that of the rest of the economy is low.

JEL: J24, J31, O30, D33

Keywords: income inequality, technological change, increasing returns to scale

*Tianxi Wang: Department of Economics, University of Essex. Email: wangt@essex.ac.uk. Tel: +44 (0)1206873480. Corresponding Author: Greg Wright: University of California, Merced, 5200 Lake Rd., Merced, 95343. Email: gwright4@ucmerced.edu. We thank various seminar participants, Gianmarco Ottaviano, Ketki Sheth and several anonymous referees.

1 Introduction

Over the last century the world has seen rapid advancements in Information and Communications Technologies (ICT) – from the early days of radio, to television, and now the internet. An important feature of ICT advancement is that it typically enables a given amount of output to be consumed or used to a greater extent, while having little impact on the actual performance of workplace tasks. For example, the invention of the radio allowed a singer's voice to be heard far beyond the walls of a theater, but of course did not allow her to sing more songs per hour. Similarly, the invention of television allowed a football match to be watched by audiences around the world without altering the way in which football is played. More recently, by expanding the reach, or "scale of operation", of workers and companies the spread of the internet has spawned a "New Economy",[1] upending traditional businesses such as retailing and air travel while at the same time spawning new occupations.[2]

When considering the impact of technological change on income the economics literature has mostly focused on the extent to which technology directly complements or substitutes for workers in performing workplace tasks. From this perspective, ICT indeed complements workers by expanding their scale of operation, and this direct effect on its own would boost worker income and attract workers into ICT-impacted occupations. However, there is an additional, countervailing general equilibrium effect: by increasing the scale of operation for *all* workers, competition between workers is heightened. For example, consider two book sellers on Amazon.com, one living in Arizona, the other in New York. Prior to the existence of Amazon these sellers served their local communities and did not compete with one another. With the advent of Amazon they now both have access to a global market, but also compete with one another as well as with many other booksellers around the world. When both of these effects are taken into account, what is the net impact of ICT on income and employment in ICT-affected occupations?

To consider this question, we formally model a worker's scale of operation. Our model starts with the observation that a worker's output can often be consumed or utilized over some range, and this range is reflected in the scale of operation. For example, a child-care worker might be able to take care of up to five children before the quality of care substantially drops; hence, his scale of operation is five children. Or while it is costly to produce a song with mass appeal, it

[1]This term seems to have originated in a *Time* magazine cover story in 1983 that discussed the transition from an industrial economy to a more technologically-oriented service economy. See "The New Economy" by Charles P. Alexander, *Time* magazine, May 30, 1983.

[2]For example, "unboxing" – in which a performer unwraps a toy in a compelling way – has become a lucrative occupation for some YouTube stars. This occupation is of course enabled by the enormous audience that the YouTube platform provides.

costs little to admit an additional person into the theater to hear it, up to the point that the theater is filled. In this case the capacity of the theater defines the scale of operation of the singer. Similarly, the capacity of a stadium defines the scale of operation for football players. The presence of fixed costs and constant marginal costs, *over some range*, is what we refer to as Increasing Returns to Scale up to some Limit (IRSL). Note that taking the limit to infinity, IRSL subsumes IRS as a special case.

We therefore model a worker's scale of operation as the limit up to which her production technology displays IRS. Furthermore, observe that ICT increases the scale of operation for some occupations – for example, singers, book sellers, and footballers – but not for other occupations, such as child-care workers. In this paper the former set of occupations are represented by a single group, *professionals*, while all of the unaffected occupations are grouped into an *alternative* occupation. In the model, workers choose between these two occupations. Importantly, the professional occupation exhibits IRSL. Formally, taking the opportunity cost of their career choice as a fixed cost (we abstract from capital and other fixed costs), each professional supplies a unique variety of services at a constant marginal cost up to a limit, denoted B. This limit represents the maximum scale of operation for a professional – for example, the capacity of the theater in which a singer performs. ICT progress is then modeled as an increase in B. The market for professional output is monopolistically competitive, such that the degree of substitutability between their varieties of services reflects the intensity of competition between them. Lastly, the quality of a professional's output increases with her human capital endowment. As a result so does her income, whereas workers in the alternative occupation, for simplicity, are assumed to receive an identical wage.

A key channel in the model is the competition effect that is alluded to above: An increase in the scale of operation, B, reduces the price of all professional services by intensifying competition between them. This creates a trade-off between this negative, general equilibrium effect and the positive, direct effect resulting from the fact that the increase in B allows each professional to sell more output. The balance of this trade-off determines the net effect on professionals' incomes and entry into the professional occupation. Specifically, while the increase in competition impacts workers identically, the ability to sell more output is of greater benefit to those who have more human capital and are therefore able to charge a higher price for their variety of services. As a result, more talented professionals – who earn higher incomes – reap greater gains from a given rise in B, and inequality within the professional occupation goes up.

Professionals that are forced into the alternative occupation due to a rise in B see their earnings fall. We show that this happens when the substitutability between professional services and the output produced by the alternative occupation is sufficiently low, leading to a sufficiently strong competition effect. The intuition for this is as follows. First, a rise in B leads to an increase in the supply of professional services, which then leads to a fall in their prices. And these prices fall more when the elasticity of demand for professional services with respect to their prices is lower. And this elasticity is lower when professional services are less substitutable with the output produced by the alternative occupation.

One counterintuitive prediction of the model is that the greater the substitutability between different professionals' services – i.e., the stronger the competition between professionals – the less likely is the marginal professional to be squeezed out due to a rise in B. The intuition is as follows. The price of each professional's services depends both on the uniqueness of her variety and the level of her human capital. Given this, the more substitutable are professional services to one another, the more their pricing will be dependent on the latter relative to the former and, as a result, the greater the price difference will be between services provided with a high, rather than low, level of human capital. In particular, the price of any *non-marginal* professional's services will be a larger multiple of the price of the *marginal* professional's services. Thus, a given decline in the price of the marginal professional's services leads to a greater decline in all other professionals' prices. When the scale of operation B rises, the average price of professional services needs to fall by a certain amount. Given this amount, the greater is the substitutability between different professionals' services, the less the marginal professional's price needs to fall, and hence the less likely she is to be squeezed out.

Finally, in order to highlight the importance of the competition effect induced by ICT we compare it to a technological change that increases productivity by reducing the marginal cost of professionals' production. The consequences for within-occupation inequality and occupational choice are quite different. Whereas an increase in the scale of operation leads to a *rise* in inequality between professionals while squeezing workers *out of* the occupation under the aforementioned condition, a reduction in the marginal cost *reduces* within-occupation inequality and always *attracts* workers into the professional occupation. This is because a reduction in the marginal cost, unlike a rise in the scale of operation, does not intensify competition between workers.

The paper proceeds as follows. Section 2 places our paper within the existing literature. Section 3 describes the model. Section 4 explores the consequences of a rise in the scale of operation for employment, income and inequality. Section 5 compares the results to the effect of a reduction in the marginal cost for professionals. Section 6 tests key model predictions. Section 7 provides concluding remarks.

2 The Literature

The general theme of our paper – that technological change may increase income inequality – fits within a large strand of literature that typically models technological change as an increase in the capital stock or as a direct increase in labor productivity. The dominant theoretical approach in this literature is the theory of skill biased technological change (SBTC hereafter).[3] This body of work focuses primarily on inequality between different skill groups, in which the effects of technological change are independent of the occupations or tasks that the workers choose.[4] In contrast, in our paper a worker is affected by technological change that increases B only if she opts into the professional occupation. Hence, our paper is closer to studies that feature task-specific technological change along with the endogeneous allocation of workers to those tasks (e.g., Cortes (2016), Autor and Handel (2013), Jung and Mercenier (2014), Acemoglu and Autor (2011a), Aum et al. (2018), Costinot and Vogel (2010) or Lee and Shin (2017)). In particular, Aum et al. (2018), and Lee and Shin (2017) show that task-specific technological change is of first-order importance for job polarization, structural change and TFP.

Relative to this tasks-based literature, our innovations are three-fold. First, our modeling of ICT is novel and captures a realistic feature of ICT, namely that it may increase a worker's scale of operation while leaving the way in which production tasks are performed unchanged. More generally, our modeling of the ICT-affected occupation subsumes as a special case the modeling of a task in the tasks-based literature. Our model reduces to the latter when professionals are homogeneous; the marginal cost of professionals' production is zero; and professional services are perfectly substitutable. Our modeling of the ICT-affected occupation is therefore richer, though our modeling of the rest of the economy is less detailed, consistent with our goal of studying the within-occupation effects of ICT.

Second, these model innovations allow us to investigate the competition effect of ICT, which has thus far not been studied much. By allowing for worker heterogeneity we show that, due to the competition effect, ICT reduces the incomes of lower-end professionals under certain conditions, despite increasing their productivity. By allowing for positive marginal costs we show that, due to the competition effect, an increase in the scale of operation leads to effects

[3]For instance, Tinbergen (1974) is an example of early work linking the demand for skill to technology; Autor et al. (2003) find that computers displace routine workplace tasks and complement cognitive-intensive, non-routine tasks; Firpo et al. (2012) find an important role for technology in generating the observed inequality pattern over the 1980s, 1990s, and 2000s; Beaudry et al. (2010) find that computer adoption increases the return to skill; and Chen et al. (2013) find that technology has increased inequality across OECD countries. Recently, Acemoglu and Autor (2011b) have extended the standard SBTC framework to endogenize the matching of skills to tasks. Eden and Gaggl (2018) explore the impact of ICT on the decline in the labor share.

[4]See e.g., Card and DiNardo (2002) and Berman et al. (1998).

that are directly opposite to those caused by a reduction in marginal costs – despite the fact that a rise in quantity is often believed to be equivalent in effect to a reduction in cost. Lastly, by allowing for a general degree of substitutability across professional services we find the counterintuitive result that the stronger the competition between professionals, the weaker the competition effect of ICT for the marginal professional. The last two results have no counterparts in the literature. With respect to the first result Lee and Shin (2017) is the only exception, as far as we know. Those authors show that if the productivity of a range of occupations simultaneously increases, in one of them workers see their relative wage fall. One difference compared to our result is that all the workers in that occupation are equally affected because they are homogeneous, whereas in our paper only the lower-end professionals definitely suffer a loss; the top-end workers may still gain.

Third, we focus on the implications of ICT for within-occupation inequality. Specifically, by incorporating worker heterogeneity we show that ICT always increases the log wage gap within affected occupations. Growth in within-occupation inequality has been an important contributor to overall inequality growth in recent years,[5] and yet has received only limited attention in the literature. Whereas tasks-based models sometimes draw an equivalence between tasks and occupations, these models assume that workers are homogenous within tasks and so abstract from within-occupation inequality.[6] Alternatively, the economy described by tasks-based models could be interpreted as representing a single occupation comprised of a continuum of tasks, as in matching models such as Sattinger (1993). However, in these models the log-wage gap does not increase when there is a rise in labor productivity within the occupation – that is, within the entire economy.

Beyond the literature that models technological change as a capital good, Garicano and Rossi-Hansberg (2014) build on Lucas Jr (1978) to consider the implications of ICT for the income distribution. Formally, they model ICT innovation as a reduction in the rate at which the marginal return to labor that is assigned to some manager falls. In this model, and in contrast to ours, ICT does not intensify competition between managers, and no one loses. Relatedly, Garicano and Rossi-Hansberg (2006, 2004) and Saint-Paul (2007) model ICT as a reduction in communication costs and consider the effect on the income distribution. In these papers knowledge production and the organization of knowledge play an important role, and are channels that are distinct from those considered in our model. Other theoretical studies on the effects of technology (not necessarily ICT) include Jones and Kim (2012) who

[5]We document that 40 percent of the growth in U.S. wage inequality over the period 1990-2010 is due to growth in within-occupation inequality; see Figure 3 and the associated discussion in Section 6.

[6]So do other studies that take the occupation or task as the unit of analysis (e.g., Autor et al. (2003) or Autor et al. (2008)).

endogenize the Pareto income distribution in a model in which technological progress augments the effects of entrepreneurs' efforts to increase productivity; and Saint-Paul (2006) who studies how productivity growth affects income inequality when consumers' utility from product variety is bounded from above.

The role of scale of operation as it relates to the return to skill has long been noted in the literature on the top end of the income distribution, i.e., the earnings of "superstars"; for example, see Rosen (1981), Rosen (1983), Gabaix and Landier (2006) and Egger and Kreickemeier (2012), and see Neal and Rosen (2000) for a summary. However, this literature is mainly concerned with income inequality for a *given* level of technology, and in particular with explaining how small differences in talent can lead to large differences in income. At the same time, it offers only an informal discussion regarding the potential impact of an increase in scale of operation. In contrast, we model the scale of operation as the limit to IRS in order to formally address this topic. In addition, we also study selection into and out of the occupations – i.e., the occupational choice margin for low-end workers (who are clearly not superstars), whereas this margin is absent in that literature.

Our model has some of the flavor of Melitz (2003),[7] though the two papers study very different issues.[8] Both papers feature monopolistic competition with CES preferences, IRS, and agent heterogeneity. However, in our paper IRS operates up to some finite limit, whereas in Melitz (2003) this limit is infinity. As a result, while an increase in the scale of operation in our paper might be regarded as parallel to an increase in the number of trading partners in Melitz (2003) (i.e., both reflect an increase in market size), the mechanism leading to reallocation is different. In Melitz (2003), it works through the factor market but it does not alter the price of any variety in the product market. In contrast, in our paper the cost of the factor is unchanged, and the increased competition works through the product market, lowering the price of all varieties. On the other hand, this product market effect is similar to Melitz and Ottaviano (2008), who present a model of monopolistic competition with quadratic preferences. However, this approach then leads to different implications for changes in market size (scale of operation). Whereas in Melitz and Ottaviano (2008) a larger market supports a greater number of varieties, in our paper an increase in the scale of operation may reduce the number of varieties, similar to an increase in the number of trading partners in Melitz (2003).

[7] More accurately, our model is in line with Melitz (2003)-style models that incorporate heterogeneity in product quality, since the heterogeneity we introduce augments the marginal value of a unit of consumption, as in those models. For instance, see Baldwin and Harrigan (2007) or Kugler and Verhoogen (2012).

[8] Specifically, our paper is concerned with the effects of technological progress on the within-occupation income distribution in a closed economy, while Melitz (2003) is focused on the relationship between exporting and aggregate productivity in an open economy.

3 The Model

The model focuses on the set of occupations for which the scale of operation is increased by ICT. To this end, we make two abstractions. First, we model all these occupations as one occupation, denoted the *professional* occupation. We assume that these occupations require particular types of human capital, which will simply be called *human capital*. Separately, *labor* is used to represent all the other attributes of workers – namely, other types of human capital and labor itself.[9] We observe that the income that workers earn is ultimately the rent that accrues to the factors of production that they contribute, though in reality it can take a variety of forms, such as wages, commissions, share of profits, etc. Second, to simplify the exposition further, we assume that all the attributes that are labeled as labor earn the same rent, denoted A. This A therefore represents the average income of the workers outside the professional occupation.

The economy is populated by a continuum of agents who choose whether to enter the professional occupation, in which ICT increases the scale of operation of workers, or to enter the alternative occupation in which ICT has no direct impact. Agent $i \in [0,1]$ is endowed with one unit of labor and h_i units of human capital. Without loss of generality, let h_i be increasing in i, that is $h_i' := \frac{dh_i}{di} \geq 0$. Labor is used for producing non-professional goods and services. To simplify exposition, all of these goods and services are bundled into one good, called the alternative good, which is used as numeraire. We assume that one unit of labor can be used to produce A units of the alternative good. Therefore, if an agent chooses the alternative occupation, her income is A.

The production of professional services requires human capital, where labor might be required to serve some auxiliary functions, e.g. as a janitor in an office. The core input is the professional's human capital. For simplicity, we assume that human capital only affects the quality of the output and abstract from its effect on quantity, which we believe is less important.[10] The impact of a professional's human capital on the quality of the services that she provides is two fold. First, some aspects of her human capital are unique, and thus so are the services that she provides, as in the canonical Krugman (1979) model. As a result, each professional provides a unique variety of professional services, indexed by her identity $i \in [0,1]$, and professionals compete under monopolistic competition. Second, professional services that are provided with a higher level of human capital are of better quality, in the sense that they

[9] This feature of an identical labor endowment and heterogeneous human capital (or ability) is also found in Lucas Jr (1978) and Monte (2011).

[10] For example, the best software engineers are the best not because they write the most lines of code per hour, but because their code generates the most sought-after software.

give greater pleasure to the consumers, as be made clear later. We abstract from the effect of human capital on output quantity by assuming that all agents have the same production function. Specifically, if an agent hires L units of labor, the output of her variety is

$$y = \left\{ \begin{array}{c} \frac{A}{c}L, \text{ if } L \leq \frac{c}{A}B \\ B, \text{ if } L > \frac{c}{A}B \end{array} \right\}, \tag{1}$$

where $c \geq 0$ is a constant. Observe that if $c = 0$, then labor is not needed for the production of the professional service and equation (1) becomes $y = B$ – i.e., each professional produces B units of output.[11] This modeling of the production technology is analogous to the modeling of task-specific technology in the tasks-based literature. We allow any $c \geq 0$. Thus, in addition to the professional's human capital, a unit of output requires c/A units of labor input before the quantity of output reaches B. Here, B represents the limit to the scale over which professionals' products can be consumed or utilized. For example, if the professional is a musician performing in a theater, then y is the number of people who can enjoy her performance and B is the capacity of the theater, which defines her scale of operation. Equation (1) implies that to admit more people into the theater, more ushers are required, until the full capacity B of the theater is reached. In this case the size of the audience is fixed at B no matter how many more ushers are employed. Similarly, if the professional is an engineer who develops software for Microsoft Windows, then y is the number of copies of the software that he sells, and B is the number of Windows users who are aware of this software, which defines his scale of operation.

Since the opportunity cost of labor is A, the marginal cost of producing one unit of services is $c/A \times A = c$. The opportunity cost of the agent's career choice (e.g., time) is A, because the alternative use of his time is to supply labor and earns income A.[12] The cost function associated with producing professional services is

$$C(y) = \left\{ \begin{array}{c} A + cy, \text{ if } y \leq B \\ \infty, \text{ if } y > B. \end{array} \right\}. \tag{2}$$

Due to the existence of fixed costs A, the average cost decreases with output y until $y \geq B$. Hence, if $B = \infty$, the production of professional services constitutes a typical instance of Increasing Returns to Scale (IRS). However, if $B < \infty$, then production of services displays

[11] Letting c converge to zero, then L in the upper branch of (1) converges to $c/A \times B$ and $y = A/c \times L$ converges to B.

[12] In the model we abstract from physical capital. If physical capital is introduced, then the cost of renting it – such as a computer for a software engineer – becomes part of the fixed cost associated with choosing to become a professional.

IRS only up to the limit B. The primary modeling innovation presented here is the introduction of this limit B of IRS. ICT increases this limit for certain occupations such as musicians or software engineers.

Agents have identical preferences. If an agent consumes s units of the alternative good and e_i units of variety i of professional services, where $i \in E$ and E is the set of varieties of professional services available on the market, then her utility is

$$\left(\mu s^{\widehat{\rho}} + \big(\int_E (h_i e_i)^\rho di\big)^{\widehat{\rho}/\rho}\right)^{1/\widehat{\rho}}.$$

where $\mu > 0$ measures the relative importance of the alternative good in the agent's utility function; $\widehat{\rho} < 1$ measures the substitutability or complementarity (as we allow $\widehat{\rho} < 0$) between the alternative good and professional services; and $\rho \in [0,1]$ measures the substitutability between one professional service and another. We assume $\widehat{\rho} < \rho$, namely that the alternative good is less substitutable for professional services than one variety of professional services is to another. In general, we think that the services or goods produced by workers in the same occupation are more homogeneous to one another relative to the services or goods produced by another occupation.[13] Note that the marginal value of agent i's services is h_i, the same as the amount of her human capital. That is, professional services provided with higher human capital deliver greater value to consumers, as noted above.

Remark: We can compare this setup with the tasks-based literature – e.g., Acemoglu and Autor (2011a), Aum, Lee and Shin (2018), and Lee and Shin (2017). In our model professionals are heterogeneous in their human capital endowments; the degree of substitutability between their services is $\rho \in [0, 1]$; and their production technology displays IRSL, as given by equation (1). In contrast, in the tasks-based literature each task is performed by homogeneous workers; their output is perfectly substitutable to one another; and each of them produces a fixed quantity that depends on the task-specific technology. This modeling of a task is a special case of our modeling of the professional occupation, in which $h_i = h$ for any i (i.e., no heterogeneity), $\rho = 1$; and $c = 0$.

Let p_i denote the price of variety i of professional services and let m denote the income of a representative agent. Then, the consumption decision that the agent faces is

[13]Some empirical evidence for this is provided in the trade literature; see e.g., Broda and Weinstein (2006) or Soderbery (2015).

$$\max_{s,\{e_i\}_{i\in E}} \Big(\mu s^{\widehat{\rho}} + \big(\int_E (h_i e_i)^\rho di\big)^{\widehat{\rho}/\rho}\Big)^{1/\widehat{\rho}},$$
$$\text{s.t. } s + \int_E p_i e_i di \le m.$$

The agent's demand for the alternative good and professional services are, respectively:

$$s = m \cdot \frac{\mu^{\frac{1}{1-\widehat{\rho}}} P^{\frac{\widehat{\rho}}{1-\widehat{\rho}}}}{1 + \mu^{\frac{1}{1-\widehat{\rho}}} P^{\frac{\widehat{\rho}}{1-\widehat{\rho}}}} \tag{3}$$
$$e_i = m \cdot f(P,\mu) \cdot h_i^{\frac{\rho}{1-\rho}} p_i^{-\frac{1}{1-\rho}}, \tag{4}$$

where P is the general price of professional services per unit of quality, defined as

$$P := \Big(\int_E (p_i/h_i)^{\rho/(\rho-1)} di\Big)^{(\rho-1)/\rho}, \tag{5}$$

and

$$f(P,\mu) := \frac{P^{\frac{\rho}{1-\rho}}}{1 + \mu^{\frac{1}{1-\widehat{\rho}}} P^{\frac{\widehat{\rho}}{1-\widehat{\rho}}}}.$$

According to equation (4), spending on a particular variety i of services

$$p_i e_i = m \frac{1}{1 + \mu^{\frac{1}{1-\widehat{\rho}}} P^{\frac{\widehat{\rho}}{1-\widehat{\rho}}}} \cdot P^{\frac{\rho}{1-\rho}} (h_i/p_i)^{\rho/(1-\rho)}.$$

For intuition on this equation, observe that, following from equation (3), $m \cdot \frac{1}{1+\mu^{\frac{1}{1-\widehat{\rho}}} P^{\frac{\widehat{\rho}}{1-\widehat{\rho}}}}$ is the agent's income spent on all available services. Following equation (4), the portion of this expenditure spent on variety i is $(h_i/p_i)^{\rho/(1-\rho)} / \int_E (h_j/p_j)^{\rho/(1-\rho)} = P^{\frac{\rho}{1-\rho}} (h_i/p_i)^{\rho/(1-\rho)}$. The spending on variety i is proportional to per-dollar quality h_i/p_i raised to the power $\rho/(1-\rho)$ because the varieties are in general not perfect substitutes for each other. In the special case in which they are – i.e. $\rho = 1$ – only the varieties with the highest per-dollar quality attract any demand.

Note that the demand for a variety is linear with the agent's income. Hence, the aggregate demand for variety i is

$$D(p_i; h_i) = M \cdot f(P,\mu) \cdot h_i^{\rho/(1-\rho)} p_i^{-1/(1-\rho)}, \tag{6}$$

where

$$M := \int_{[0,1]} m_j dj \tag{7}$$

is aggregate income. Note that $D'_h > 0$ – that is, given the price, the demand for a higher-quality variety is greater because consumers derive greater value from it.

If agent i chooses to enter the alternative occupations, she gets A. If the agent chooses to enter the professional occupation and produce her variety of services at marginal cost c, the demand for her services will be given by (6), where she takes the aggregate variables P and M as given. She then sets the price of her services by solving the following decision problem:

$$m(h_i) = \max_{p_i}(p_i - c)D(p_i; h_i), \text{ s.t. } D(p_i; h_i) \leq B. \tag{8}$$

The agent chooses to enter the professional occupation rather than the alternative occupation only if

$$m(h_i) \geq A. \tag{9}$$

From the envelope theorem and (8), $m'(h) > 0$. There thus exists a threshold $k \in [0,1]$ such that agent i chooses to enter the professional occupation if and only if $i \geq k$, where k is pinned down by

$$m(h_k) = A. \tag{10}$$

Hence the set of available professional services is $E = [k, 1]$. It follows that the general price for professional services, from (5), is given by

$$P = \Big(\int_k^1 (p_i/h_i)^{\rho/(\rho-1)} di\Big)^{(\rho-1)/\rho}. \tag{11}$$

and if $i < k$ agent i earns income A, while if $i \geq k$ agent i earns $m(h_i)$, hence aggregate income is

$$M{=}kA + \int_k^1 m(h_i)di. \tag{12}$$

Definition 1. *A profile (P, k, M) forms a competitive equilibrium if*

(i) P is given by (11), where p_i solves (8);

(ii) agent i chooses to enter the alternative occupation if and only if $i < k$ where k is determined by (10);

(iii) Aggregate income is given by (12).[14]

4 An Increase in the Scale of Operation in the Professional Occupation

In this section we prove the existence of a unique equilibrium, find the equilibrium income of each agent, and then consider the effect of an increase in B, which represents ICT progress.

4.1 The existence and characterization of a unique equilibrium

We first focus on the case in which the capacity constraint, $D(p;h) \leq B$, is binding for all agents who choose to be a professional. This is equivalent to requiring that $B < A\rho/(c(1-\rho))$, as we show in Subsection 4.3, where we also show that the insights derived from this case can then be applied straightforwardly to the case in which the capacity constraint is not binding for a subset of professionals. Of course, if it is not binding for any agents then an increase in B will have no effect.

With $D(p_i;h_i)$ given by (6), the binding capacity constraint, $D(p_i;h_i)= B$, implies that the price of variety i is:

$$p_i = \Big(\frac{Mf(P,\mu)}{B}\Big)^{1-\rho} h_i^{\rho}. \tag{13}$$

Thus, an agent with higher human capital charges a higher price for her services because they deliver greater value to consumers. In fact, the price is proportional to the marginal value raised to power $\rho < 1$ – that is, h_i^{ρ}. It follows that for any $i \geq k$, $p_i/p_k = h_i^{\rho}/h_k^{\rho}$. Hence, $p_i B/\int_k^1 (p_j B dj) = h_i^{\rho}/\int_k^1 h_j^{\rho} dj$, that is, aggregate spending on variety i is a fraction, h_i^{ρ}/H_k^{ρ}, of the aggregate spending on all varieties of services, where for any $x \in [0,1]$, we define

$$H_x := \Big\{\int_x^1 h_i^{\rho} di\Big\}^{\frac{1}{\rho}}. \tag{14}$$

Agent k, the marginal professional, obtains profit A because she is indifferent between the two occupational choices. That is, $(p_k - c)\times B = A$, or

$$p_k = A/B + c. \tag{15}$$

It follows that $p_i = (A/B + c) \times h_i^{\rho}/h_k^{\rho}$ and the income of agents $i \geq k$ is $m_i = (p_i - c)B = (A + Bc)h_i^{\rho}/h_k^{\rho} - Bc$. We know that agents $i < k$ choose to enter the alternative occupation

[14] We skip the clearing of the alternative good market, which pins down the fraction of labor used for producing the good, a variable that is not very interesting in the context of this paper.

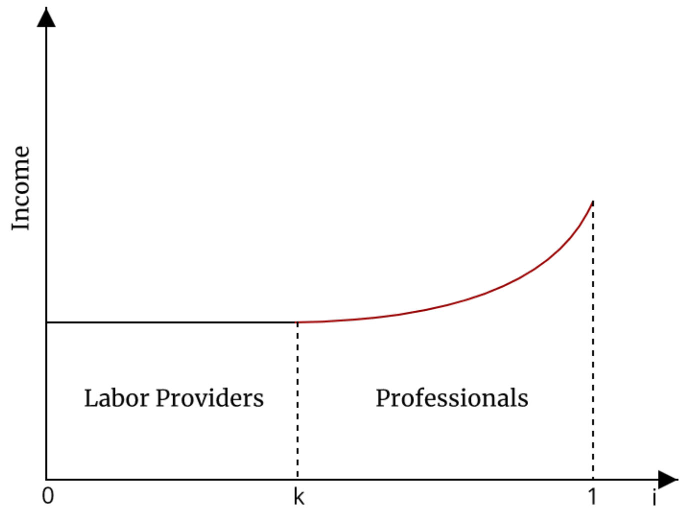

Figure 1: The Equilibrium Income Distribution

and earn $m_i = A$. Putting these together, the equilibrium income distribution is:

$$m_i = \begin{cases} A \text{ if } i < k \\ (Bc + A)\frac{h_i^\rho}{h_k^\rho} - Bc \text{ if } i \geq k \end{cases} \tag{16}$$

This income distribution is illustrated in Figure 1.[15]

We find that k is determined by the following market clearing condition for the alternative good, where

$$k_0 := \frac{Bc}{A + Bc}. \tag{17}$$

Proposition 2. *The identity of the marginal professional, k, is determined by*

$$\mu^{\frac{1}{1-\widehat{\rho}}} H_k^{\frac{\rho-\widehat{\rho}}{1-\widehat{\rho}}} h_k^{\frac{-\rho}{1-\widehat{\rho}}} \times (A/B + c)^{\frac{\widehat{\rho}}{1-\widehat{\rho}}} = k - k_0. \tag{18}$$

The equation has a unique root for k over $(k_0, 1)$. Hence the equilibrium uniquely exists.

Proof. We relegate the proof to B.1. ∎

[15]The figure is based on the assumption that h_i is a convex function of i so that m_i, though a concave function of h_i, is convex in i. Roughly, the assumption is that within a typical talent distribution, there are a small number of people at the top who are much more talented than the rest – a view that seems consistent with the evidence.

This equation is essentially a market clearing condition for the alternative good. First, the term on the right hand side represents aggregate supply in units of the marginal professional's revenue, $A + Bc$. If the marginal professional is k, then a mass $1 - k$ of agents choose the professional occupation and in aggregate they use $(1-k)\,B \times c/A$ units of labor. Hence, a mass $k - (1-k)\,Bc/A$ of agents produces the alternative good. In aggregate they supply $[k - (1-k)\,Bc/A] \times A = (k - k_0) \times (A + Bc)$. Second, the term on the left hand side of (18) represents the aggregate spending on the alternative good in units of the marginal professional's revenue. This can be clearly seen for the Cobb-Douglas case, in which $\widehat{\rho} = 0$. In this case, this term simplifies to $\mu H_k^{\rho}/h_k^{\rho}$. Measured in units of the marginal professional's revenue, the spending on his service (the agent's revenue) is 1. We saw that this spending is a fraction h_k^{ρ}/H_k^{ρ} of the aggregate spending on all services. The aggregate spending on services is therefore H_k^{ρ}/h_k^{ρ}, and then μ times this term gives the aggregate spending on the alternative good in the Cobb-Douglas case, where the ratio of the spending on the alternative good to that on services is always μ, independent of the price of services.[16]

For the non-Cobb-Douglas case, the effect of services prices on aggregate spending on the alternative good is summarized by the term $(A/B + c)^{\frac{\widehat{\rho}}{1-\widehat{\rho}}} = (p_k)^{\frac{\widehat{\rho}}{1-\widehat{\rho}}}$. Here the effect depends on the sign of $\widehat{\rho}$. This is because a change in service prices generates two standard, and conflicting, effects on the demand for the alternative good – namely, the substitution and income effects. If $\widehat{\rho} < 0$, the income effect dominates. Hence, when professional services become cheaper, reflected in a smaller p_k, spending on the alternative good goes up, and the opposite occurs when $\widehat{\rho} > 0$.

The equilibrium is unique because the two sides of the alternative market vary monotonically with k, but in opposite directions. On the one hand, the aggregate supply of the alternative good increases with k due to the fact that the larger is k, the more agents there are that produce the alternative good. On the other hand, aggregate spending on the good decreases with k and goes to zero as k goes to 1.[17] Intuitively, this is because if an agent with little human capital chooses to become a professional, it must mean that the economy is sufficiently rich, and therefore spends a significant amount on professional services. Put differently, if the threshold k of human capital for entering the professional occupation is rising, that is because the economy is getting poorer, which means that aggregate spending on the alternative good is shrinking too. In the extreme case, if the economy can support only the agent with the greatest

[16] In the Cobb-Douglas case, each agent spends a fraction of $\frac{\mu}{1+\mu}$ of his income on the subsistence good and $\frac{1}{1+\mu}$ on entertainment services. Hence, aggregate spending on the former good is μ times spending on the latter good.

[17] Since $\rho - \widehat{\rho} > 0$, $H_k^{\frac{\rho-\widehat{\rho}}{1-\widehat{\rho}}}$ increases with $H_k = \{\int_k^1 h_i^{\rho} di\}^{\frac{1}{\rho}}$, which decreases with k. Since $\rho > 0$, $h_k^{\frac{-\rho}{1-\widehat{\rho}}}$ decreases with h_k which, by assumption, increases with k. Moreover, $H_1 = 0$. Hence the term on the left hand side equals 0 at $k = 1$.

human capital as a professional – i.e., $k = 1$ – then it must be extremely poor and aggregate income must be approaching zero.

4.2 ICT, Inequality and Occupational Choice

In this section we explore the impact of a rise in the scale of operation B on income inequality within the professional occupation as well as agents' occupational choice. Intuitively, an increase in B affects the incomes of all professionals, given by $(p - c) B$, in two ways. First, an increase in B expands the capacity of all professionals and enables them to sell their services to more consumers, a positive effect. Second, since this expansion is of equal magnitude for all professionals, it necessarily leads to stronger competition between them, which should cause the price of all professional services to fall, a negative effect. However, observe that if an agent is able to charge a higher price, she gains more from the enlargement of capacity. Since agents with higher levels of human capital charge higher prices and earn greater income, this suggests that the more an agent is earning presently, the more she gains from the positive effect of a rise in B. As a result, inequality between professionals should increase. This intuition is confirmed by the following proposition.

Proposition 3. *For $i > k$, $\frac{d \log m_i}{dB}$ strictly increases with m_i – namely the rate of change in the income of professionals induced by a rise in B is positively correlated with their present income.*

Proof. We relegate the proof to B.2. ∎

Following from this proposition, if one regresses the percentage change of a professional's income on the product of his present income and some proxy for an increase in the scale of operation, then one should expect the coefficient to be positive. Observe that while the changes to the highest-earning professionals follows the pattern of Proposition 3, their relative positions may not be affected much by a rise in B. This is because if both $h_i >> h_k$ and $h_j >> h_k$, then by (16) $m_i/m_j \approx h_i^\rho / h_j^\rho$, which is independent of B. Given that this result is only focused on a comparison of the highest-earning professionals, it is not in conflict with the observation that the top 1 percent, or top 0.1 percent, of U.S. earners have reaped an increasing proportion of income in recent years.

In the above discussion, we expect that a rise in B should cause the prices of all professional services to fall due to the competition effect. This is confirmed by the following proposition.

Proposition 4. *For $i \geq k$, $\frac{dp_i}{dB} < 0$. That is, the price of each variety of services falls when the scale of operation, B, rises.*

Proof. We relegate the proof to B.3. ∎

Having examined the implications of a rise in B for within-occupation inequality, we now turn to its effect on agents' occupational choice, as reflected in a movement in the cutoff k. Studying this effect is not only interesting in itself, but also has implications for the net effect for individual professionals' incomes. For example, if an increase in scale of operation causes the cutoff k to fall, then that means that the marginal professional is not indifferent between the two occupational options any more, but instead strictly prefers becoming a professional. That is, the marginal professional is better off. Since the percentage change in other professionals' incomes is even greater – by Proposition 3 – all professionals reap a net gain from the rise in B. If, instead, the rise causes the cutoff to go up, from k to some $k' > k$, then agents $i \in (k, k']$ who previously received income $m_i > A$ now choose the alternative occupation and thus receive income A. That is, these agents are worse off.

This discussion indicates that the movement in the cutoff k depends on how the two aforementioned effects of a rise in B balance out for the marginal professional. If the positive effect due to the expansion in market size dominates, then the cutoff k falls. In contrast, if the negative competition effect dominates, the marginal professional exits and the cutoff k rises. The following proposition determines the sufficient and necessary condition for either scenario to arise in equilibrium.

Proposition 5. $\frac{dk}{dB} > 0$ *if and only if*

$$\widehat{\rho} \leq k_0; \;\; \textit{or } \mu < \eta \left(\frac{h_{k_0/\widehat{\rho}}}{H_{k_0/\widehat{\rho}}} \right)^{\rho}, \tag{19}$$

where η *is a constant independent of* ρ *and* $\eta = 0$ *if* $k_0 = 0$.

According to Proposition 5, under condition (19), with a rise in B the negative competition effect dominates the positive market-size effect for the lower-end professionals and, as a result, they are squeezed out of the professional occupation and enter the alternative occupation. Condition (19) requires either that the substitutability between professional services and the alternative good $\widehat{\rho}$ is not too large or that the relative importance of the alternative good to professional services, μ, is not too high. To understand why a rise in B lowers the prices of services by such a large amount, observe that a fall in the services price shifts agents' demand away from the alternative good toward the professional services. This channel absorbs the increase in the supply of services caused by the rise in B. As a result, the services price will fall by a large amount if this demand side channel is weak, which is the case if $\widehat{\rho}$ or μ is small enough.

If $\widehat{\rho}$ is small – that is, if professional services are not very substitutable with the alternative good – then consuming more of the former diminishes very little of agents' valuation of the latter; indeed, they value the alternative good even more if $\widehat{\rho} < 0$ – i.e., if professional services are complementary to the alternative good. In this scenario, a rise in the supply of professional services is absorbed very little by a reduction in the demand for the alternative good. The same is true if $\widehat{\rho}$ is big but μ is small because, although now professional services are highly substitutable with the alternative good, one unit less of the alternative good will be absorbed by – that is, substitute for – only μ units of professional services.

The value of ρ has also an impact on the sign of dk/dB. For any given x, h_x^ρ/H_x^ρ decreases with ρ because $H_x^\rho/h_x^\rho = \int_x^1 (h_i/h_x)^\rho \, di$ increases with ρ, since $h_i/h_x > 1$. Hence, the greater is the value of ρ, the less likely that the condition $\mu < \eta \left(h_{k_0/\widehat{\rho}}/H_{k_0/\widehat{\rho}}\right)^\rho$ holds true, and thus the less likely $dk/dB > 0$. That is, the stronger the competition between professionals, the less likely the marginal professional is squeezed out due to a rise in the scale of operation, a counter-intuitive result. The intuition is as follows. Recall that for any professional i, the price she charges is $(h_i/h_k)^\rho$ times the price charged by the marginal professional, that is: $p_i = (h_i/h_k)^\rho \times p_k$. For any $i > k$, $(h_i/h_k)^\rho$ increases with ρ. Therefore, the price charged by other professionals will be a greater multiple of the marginal professional's price when varieties of professional services are more substitutable – i.e., when ρ is larger. Put differently, if ρ is greater, a given decline in p_k leads p_i to fall by a larger amount for all $i > k$. When the scale of operation B rises, the average price of professional services needs to fall by a certain amount that depends mainly on $\widehat{\rho}$. Given this amount, the greater is ρ, the smaller the scale by which the marginal professional's price needs to fall and hence the less likely she is to be squeezed out. This effect seems counter-intuitive, but it is related to the sign of $\partial^2 k/\partial B \partial \rho$ and should not be confused with the result that, given B, the stronger the competition between professionals, the smaller the number of workers staying the occupation – that is, $\partial k/\partial \rho > 0$, which can be shown to hold using equation (18).

As argued above, one corollary of an increase in the cutoff k is that lower end professionals suffer a net loss. This is confirmed in the following corollary.

Corollary 6. *Under condition (19) there exists $\widehat{k} > k$ such that $dm_i/dB < 0$ for $i \in (k, \widehat{k})$ – i.e., lower-end professionals lose as the result of an increase in the scale of operation.*

Proof. We relegate the proof to B.3. ∎

Proposition 5, together with its corollary, highlights a unique feature of technological progress that increases the scale of operation, B. Specifically, an increase in scale of operation complements workers by increasing the size of the market for their output, much like a factor augmenting technology and, as a result, would be expected to increase the incomes of all the workers and induce entry into the affected occupations. However, there is a countervailing force, which is the competition effect. Under condition (19), this countervailing force dominates, lowering some workers' incomes[18] and inducing exit from the affected occupations.

While lower-end professionals clearly lose under condition (19), the top end could still gain. As noted, the gain due to the market size expansion is larger if the professional charges a higher price p, which is proportional to his human capital endowment raised to the power ρ. If a professional's human capital endowment is high enough then the gains will outweigh the losses due to fiercer competition, and the professional will reap a net benefit due to the increase in B. To find a condition under which this happens, let

$$\Omega(\rho) := \max_{x\in[k_0,1]} \frac{\rho \cdot h'(x)/h(x)}{1+\rho \cdot h'(x)/h(x)\cdot(x-k_0)},$$

and if $\Omega(\rho)\cdot A/(A+Bc) < 1$ let

$$\xi := [\frac{1}{1-\Omega(\rho)\cdot A/(A+Bc)}]^{\frac{1}{1-\widehat{\rho}}}.$$

Furthermore, let $f(k_0, y)$ denote the unique solution for $t \in [k_0, 1]$ in

$$t - k_0 = y(1-t)^{\frac{\rho-\widehat{\rho}}{\rho(1-\widehat{\rho})}},$$

and let

$$D := \mu^{\frac{1}{1-\widehat{\rho}}}(A/B+c)^{\frac{\widehat{\rho}}{1-\widehat{\rho}}}.$$

The following proposition provides conditions under which the top professional reaps a net gain from an increase in B.

Proposition 7. *Assume that* $\Omega(\rho)\cdot A/(A+Bc) < 1$ *and* $\widehat{\rho} \geq 0$. *If* $h_1 > 1$ *and* $h_1 \geq \xi \cdot h(f(k_0, D\cdot \xi))$, *then* $dm_1/dB > 0$.

Proof. We relegate the proof to B.6. ∎

[18]One caveat is that these workers' welfare might still be higher, despite their income, as measured relative to the constant real good (i.e. the alternative good), being reduced, because all the professional services are cheaper (Proposition 4).

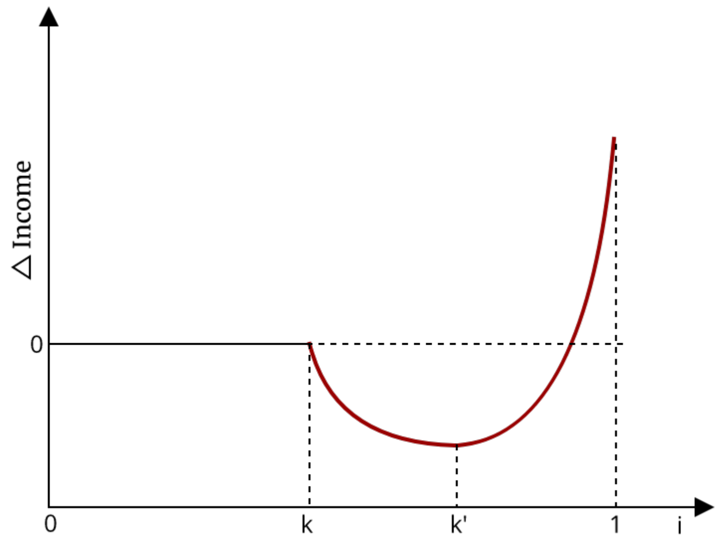

Figure 2: Income growth due to an expansion in B.

Note that this proposition is concerned with only two points along the function $h(i)$, namely at $i = 1$ and $i = f(k_0, D \cdot \xi)$, and hence it can be satisfied by any distribution of human capital in which $h(1)$ is sufficiently large. When this condition holds, the top professionals gain in net from an increase in B, while under condition (19), professionals at the bottom of the distribution suffers a net loss. Hence, under the conditions of Proposition 7 and condition (19), an increase in B leads to a U-shaped change in income across agents, as illustrated in Figure 2.

4.3 When the Capacity Constraint Is Non-Binding for Some Professionals

Thus far we have considered the case in which the capacity constraint, $D(p;h) \leq B$, is binding for all professionals. If the capacity constraint is non-binding for some professionals, then these professionals' human capital will lie at the lower end of the distribution. The demand for a professional's services, by (6), is proportional to $h_i^{\rho/(1-\rho)}$. Thus, the profit-maximizing output in the absence of the capacity constraint increases with h_i. As a result, if it is binding for agent i then it is binding for all the agents $i' \geq i$, and if it is not binding for agent i, then neither is it for any agent $i' \leq i$. Thus, if and only if the capacity constraint is binding for the marginal professional k, will it be binding for all professionals. Since the professionals' problem is given by (8), in the absence of a capacity constraint, the optimal price is c/ρ. The constraint is binding for agent k if he cannot reach this price by supplying enough output, namely if the price pinned down by the binding capacity constraint, p_k, is above c/ρ. According to (15), $p_k = A/B + c$ in the equilibrium in which the constraint is binding for the marginal professional. Hence, the condition under which the capacity constraint is binding for all professionals is $A/B + c > c/\rho$, or equivalently,

$$\frac{\rho}{1-\rho}\frac{A}{c} > B. \tag{20}$$

Observe that the smaller the marginal cost of producing professional services (c), or the stronger the competition between different varieties of the services (i.e., a bigger ρ), the more likely it is that this condition holds. In particular, if $c = 0$, then it always holds.

If the condition does not hold then the capacity constraint is binding for some share of professionals and non-binding for the remainder. In this case, the argument above implies that there exists $i^* \in (k, 1)$ such that it is non-binding for $i < i^*$ and binding for $i > i^*$. Obviously, an increase in B makes the capacity constraint non-binding for more professionals – that is, $di^*/dB > 0$.

Here we consider how the results in the preceding subsection change in the case in which the capacity constraint is binding for only some subset of professionals. First, Proposition 2 still holds and the unique equilibrium still exists. It is driven by the same economic forces as before. If too many agents choose to provide labor and produce the alternative good, then the professional services will be expensive, which will induce further entry. Conversely, if too few agents provide labor there will be entry into production of the alternative good.

Second, Proposition 3 still holds, that is, an increases in B raises inequality within the professional occupation.

Proposition 8. *Suppose there exists $i^* \in (k, 1)$ such that the capacity constraint is non-binding for $i < i^*$ and strictly binding for $i > i^*$. Then for $i > k$, $\frac{d\log m_i}{dB}$ increases with m_i and this increase is strict for $i > i^*$.*

Proof. We relegate the proof to B.7. ∎

This proposition is driven by the same intuition as before: an enlargement in market size delivers greater benefits to professionals who charge a higher price for their services – that is, those who currently have higher incomes.

Third, Proposition 4 also holds. A rise in B intensifies competition between different varieties of professional services, causing their prices to fall.

Proposition 9. *For $i \geq k$, $\frac{dp_i}{dB} \leq 0$ and the inequality is strict for $i > i^*$.*

Proof. We relegate the proof to B.8. ∎

Fourth, we expect Proposition 5 to hold under a condition that is less strict relative to (19). That is because the marginal professional, now with a non-binding capacity constraint, gains nothing from an increase in B, while she is still subject to the negative competition effect. Without any offsetting positive effect, the marginal professional is now more likely to be pushed out of the occupation.

Finally, even if an increase in B leads to a rise in the cutoff k such that bottom-end professionals lose (Corollary 6), it still delivers a net gain to the top professionals when their human capital is high enough. This is because the benefits from the enlargement in market size are proportional to $p - c$ and the price is proportional to the human capital level h raised to the power ρ (imagine $h_1 \to \infty$ and hence $p_1 \to \infty$). Therefore, Proposition 7 holds, although the exact conditions that describe what the term "high enough" means will be different.

5 A Comparison to a Reduction in Marginal Cost

As we have argued throughout, and as highlighted by Corollary 6 in Subsection 4.2, this particular feature of ICT is special in that although it increases professionals' productivity by enabling each of them to sell to more buyers, it ultimately squeezes lower-end professionals out of the occupation and generates a net loss for the next layer of professionals under condition (19). This result arises because of the competition effect that ICT induces: by enlarging the size of the market for all professionals, it intensifies competition between them. Our modeling of ICT as an increase in B captures this competition effect. The results would be very different if the impact of ICT was modeled as a reduction in the marginal cost c of professionals' production, due to the absence of the competition effect, as we show below. Observe that we therefore present a case in which a rise in quantity B generates very different effects than a reduction in cost c, despite the fact that these two changes are often regarded as equivalent to each other.

Below we present a comparative statics analysis with respect to a decline in marginal cost, c. For the benefit of exposition, in this section we focus on the case in which condition

$$\frac{\rho}{1-\rho}\frac{A}{c} > B$$

holds and hence the capacity constraint is binding for all professionals. A reduction in c makes the condition more likely hold. The effect of a reduction in c on the cutoff k and inequality within the professional occupation is as follows.

Proposition 10. $\frac{dk}{d(-c)} < 0$ *and* $\frac{d \log m_i}{d(-c)}$ *is negatively correlated with* m_i.

Proof. We relegate the proof to B.9. ∎

According to this proposition, the effect of a reduction in the marginal cost, c, is the opposite of that due to an increase in the scale of operation B under condition (19). To begin with, a reduction in marginal costs decreases inequality within the professional occupation, whereas an increase in scale of operation increases inequality. Additionally, a reduction in c always lowers the cutoff k, drawing more workers into the professional occupation, whereas a rise in B under condition (19) raises k, squeezing workers out of the professional occupation. As a corollary, a reduction in marginal costs leads to gains for bottom-end professionals, whereas those workers lose due to an increase in B. Our model therefore demonstrates that technological changes that increase output while holding marginal costs fixed can generate effects that are entirely opposite to those generated by a reduction in marginal costs at a given level of output. The competition effect is the key channel driving this difference in outcomes.

6 Empirical Evidence

In this section we exploit U.S. occupational data over three decades to explore the key predictions of the model. Specifically, we test Propositions 2 and 4 and Corollary 1 from the model, exploiting the advent of the internet as a natural experiment. To reiterate, each of these makes a prediction about the effect of a rise in the IRS limit. Proposition 2 states that the log wage gap between workers within affected occupations will rise; Proposition 4 implies a fall in employment in affected occupations; and Corollary 1 indicates that some workers will see falling incomes.

We focus on the U.S. in order to exploit detailed annual data on workers' hours and earnings. Specifically, we use data on wages and employment within U.S. occupations from the U.S. Current Population Survey (CPS) over the years 1985 to 2010.[19] We deal with top-coding in the manner described by Bakija et al. (2010), though our results are also robust to excluding top earners. Consistent with the literature, we restrict the sample to full-time workers between 18 and 65. Our unit of interest is the occupation, and we adopt a consistent definition of occupations across datasets using the definitions from Autor and Dorn (2013).

It would be difficult to accurately measure the scale of operation for each occupation. Here, we are instead satisfied with a measure the change of which reflects the change in the scale of operation induced by the expansion of the internet. Specifically, we measure the value of goods and services sold over the internet by workers within an occupation, and denote it as B^{int}. We define the measure in the following way:

[19] The data were obtained from IPUMS (see Ruggles et al. (2010)).

Top 10		Bottom 10	
1	Financial services sales occupations	**327**	Legislators
2	Motion picture projectionists	**328**	Clergy and religious workers
3	Cabinetmakers and bench carpenters	**329**	Inspectors of agricultural products
4	Editors and reporters	**330**	Welfare service aides
5	Furniture and wood finishers	**331**	Postmasters and mail superintendents
6	Typesetters and Compositors	**332**	Meter readers
7	Other financial specialists	**333**	Mail and paper handlers
8	Broadcast equipment operators	**334**	Hotel clerks
9	Computer Software Developers	**335**	Judges
10	Actors, directors, producers	**336**	Sheriffs, bailiffs, correctional institution officers

Table 1: Top and Bottom 10 Occupations by Exposure to Internet Sales

$$B_{it}^{int} = \sum_{j} \left(IntShr_{jt} \times OccShr_{ij,1990} \right) \tag{21}$$

where $t \in \{1990, 2000, 2010\}$; $IntShr_{jt}$ is the share of industry j sales in year t that was made over the internet and is set equal to 0 in 1990; and $OccShr_{ij,1990}$ is the share of occupation i's total hours employed in industry j in 1990.[20] Thus, the latter term reflects the importance of each industry, in terms of labor hours, to each occupation in a period in which the internet was absent, where we use a pre-period occupational structure in order to avoid incorporating effects due to endogenous changes in the composition of occupations caused by the internet. The former term, $IntShr_{jt}$, then captures the extent to which firms within each industry sell their output over the internet.[21] Table 1 lists the top 10 (left column) and bottom 10 (right column) occupations in terms of their exposure to internet sales according to the 2010 measure. These lists align intuitively with the types of output that can, and cannot, be sold over the internet.

To highlight the role that occupations have played in driving the pattern of inequality growth over recent decades, Figure 3 decomposes the total rise in income inequality between 1990 and 2010 into within- and between-occupation components (see Appendix A for the details of the decomposition). Here we see that both have been important. More specifically, approximately

[20]Internet sales by industry come from Census' E-Stats database, available at http://www.census.gov/econ/estats/.

[21]Of course, the measure may not perfectly capture the extent to which occupational services are linked to internet sales. For instance, even within an industry that sells a substantial amount over the internet, some occupations may be specialized in brick-and-mortar sales, while others are focused on internet sales. Furthermore, our analysis below will focus in part on the implications for wages, but the elasticity of occupational wages to internet sales may vary across occupations for many reasons, from which we abstract.

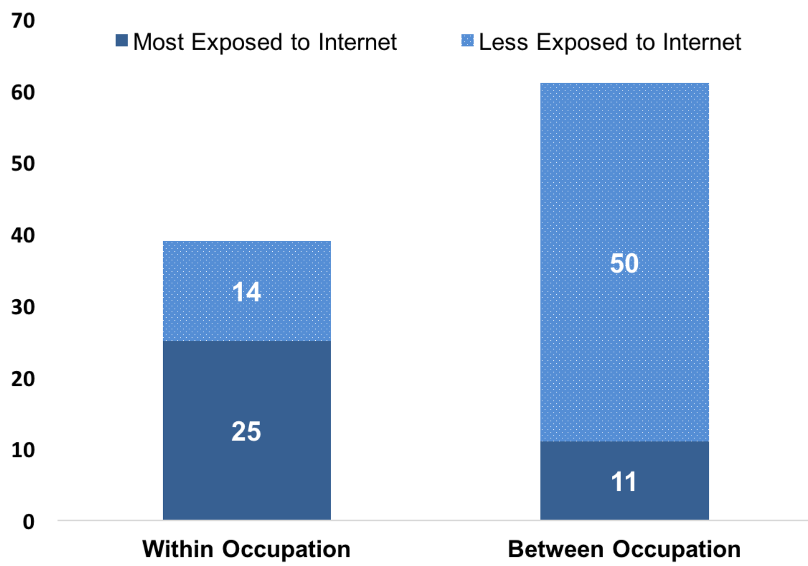

Figure 3: Contributions to Growth in U.S. Wage Inequality, 1990-2010. "Most Exposed to Internet" is defined as the top 10 percent of occupations according to our measure described above, and "Least Exposed to Internet" are all others.

40 percent of the rise in aggregate wage inequality over the period 1990-2010 occurred *within* occupations. Incorporating differences in internet exposure across occupations using measure (21) we see that nearly two-thirds of the within-occupation rise in inequality is due to the top ten percent of occupations that were most exposed to rising sales over the internet. Thus, rising within-occupation inequality has been important and, at the same time, can be largely explained by the subset of occupations that has been most impacted by the internet. In fact, the first two facts imply that over a quarter of the recent rise in *aggregate* wage inequality can be explained by these types of occupations (though of course not all of this will be explained by a rise in B).

Next, we exploit the natural experiment generated by the growth in use of the internet in the mid-1990s, which differentially exposed workers to a rise in market access for their services over subsequent years. Due to the internet, some workers, such as actors and software engineers, now face a potential market size that is many millions of consumers larger than it was in 1990, while others, such as bus drivers, face approximately the same set of potential consumers as they did in 1990.

Corollary 1 indicates that the low-end workers lose with a rise in B. Because we set $IntShr_{jt} = 0$ in 1990, the most exposed occupations – i.e., those with the largest values of $B^{int}_{i,2010}$ – are the ones who see the largest increases in B due to the internet. As a result, lower-

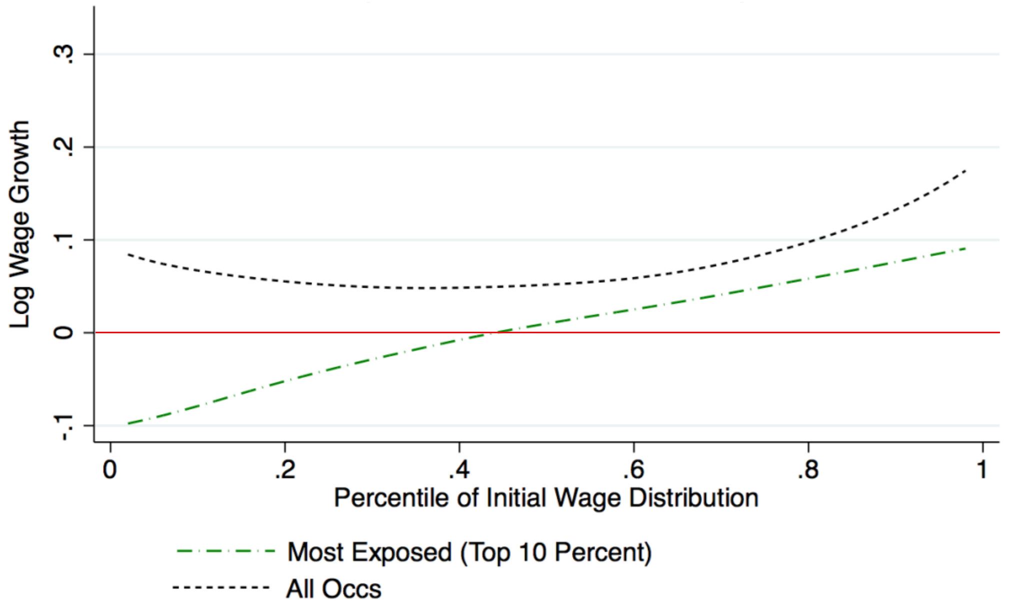

Figure 4: Occupation Exposure to the Internet and Log Wage Growth, 1990-2010

end workers in these occupations should be more likely to experience falling wages relative to those in less exposed occupations. In Figure 4 we see that workers in the most exposed occupations indeed see declining wages at the bottom of the distribution relative to the economy-wide average. We note that in this Figure the wage values have been cleaned of variation due to age, age squared, sex and level of education in a "first stage" regression; in other words, we effectively control for changes in the demographic composition of workers within occupations. We also clean the wage of industry variation (i.e., include industry fixed effects in the first stage), thereby controlling for differences across industries in the evolution of the wage structure.

Proposition 2 predicts that for any two percentiles of the wage distribution, $i > j$, $\frac{d \log w_i}{dB} - \frac{d \log w_j}{dB} > 0$. It follows that $\frac{d(\log w_i - \log w_j)}{dB} > 0$. For instance, $\frac{d(\log w_{90} - \log w_{10})}{dB} > 0$ and $\frac{d(\log w_{90} - \log w_{50})}{dB} > 0$; that is, the log wage gap between 90th percentile and 10th percentile, or 90th percentile and 50th percentile, are predicted to be increasing with B. We test Proposition 2 by estimating the following specification:

$$\triangle WageGap_{i,t:t-1} = c + \beta_1 \left(\triangle B^{int}_{i,t:t-1}\right) + \beta_2 CompUse_{i,1989} + WageGap_{i,80-90} + \epsilon_{it} \qquad (22)$$

where we follow recent convention by defining the change in wage inequality, $\triangle WageGap_{i,t:t-1}$, as the change in the gap between the 90th and 10th percentiles of the log wage distribution for

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Δ 90-10 Log Wage Gap			Δ 90-50 Log Wage Gap			Δ Log Employment		
$\Delta InternetSales$	0.354***	0.315*	0.270*	0.142**	0.113**	0.091	-0.023***	-0.021**	-0.018**
	(0.155)	(0.167)	(0.141)	(0.062)	(0.045)	(0.073)	(0.007)	(0.010)	(0.008)
$CompUse_{1989}$		0.064***	0.063***		0.026**	0.024*		0.044*	0.051*
		(0.018)	(0.018)		(0.013)	(0.014)		(0.022)	(0.028)
$WageGap^{90-10}_{1980-1990}$			-0.225***						-0.356***
			(0.084)						(0.061)
$WageGap^{90-50}_{1980-1990}$						-0.073			
						(0.101)			
Observations	642	584	584	642	584	584	642	584	584

Notes: Dependent variables are the stacked change in the 90-10 log wage gap, 90-50 log wage gap and log employment over 1990-2010. In a first-stage regression, wage variation is cleaned of variation in age, age squared, sex, education, and industry. Exposure to internet sales at the occupation level is defined as defined by equation (21). Computer use in 1989 is obtained from the computer use supplement of the 1989 Current Population Survey. The wage gap controls are the pre-period changes in the wage gap (between 1980 and 1990) for the 90-10 and 90-50 gaps, respectively. Standard errors are in parentheses and are clustered at the occupation level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: The Rise in Internet Sales and Within-Occupation Inequality and Employment, 1990-2010

each occupation i or, alternatively, the 90-50 log wage gap, where we stack changes over the periods 1990 to 2000 and 2000 to 2010. Again, the wage variation we exploit has been cleaned in a first stage, as described above. In addition, in some specifications we control for differential pre-period trends (1980 to 1990) in the wage gap, denoted $WageGap_{i,80-90}$. Standard errors are clustered at the occupation level.

In our strictest specifications we also control for the differential use of computers across occupations in 1989 ($CompUse_{i,1989}$), several years prior to the spread of the internet. Specifically, we control for the share of hours worked in an occupation by workers who use a computer in order to address the possibility that computer use is a key omitted variable in the specification.[22] In other words, the relationship between the change in internet sales within an occupation and rising inequality may be due to the direct impact of increased computer use on both wage inequality and rising internet sales. By controlling for computer use in 1989 we absorb variation in the prevalence of computer use across occupations that is unrelated to (future) internet sales.[23]

Proposition 2 predicts that β_1 should be positive. Noting that the results are suggestive, and

[22]We obtain these data from the Current Population Survey computer use supplement, 1989.

[23]The shortcoming of this control is that it likely does not perfectly predict future computer use in an occupation, and the unexplained portion of future computer use may be correlated with both rising wage inequality and internet sales. At the same time we note that some of this additional variation will likely be absorbed by the pre-trends in the wage gap.

not definitive, Table 2 presents the estimates.[24] In columns (1)-(6) we see that rising internet sales are associated with a widening wage gap, both across the 90-10 distribution as well as the 90-50 distribution. Aside from a lack of statistical significance in column (6), these effects hold across specifications, indicating that the effects are not solely driven by predicted computer use or pre-trends in the wage distribution. The economic magnitudes are also important: the estimates imply that the increase in occupational exposure to internet sales over the period explains 39 percent of the rise in the 90-10 wage gap and 26 percent of the rise in the 90-50 gap.[25]

With respect to Proposition 4 – which predicts that employment will fall in affected occupations in response to a rise in B^{int} – we first note that employment growth among the top ten percent of occupations most linked to internet sales was 7 percent over the period, rising from about 7 to 7.5 million workers, significantly slower than the 36 percent average employment growth associated with other occupations. We then estimate a specification identical to (22) except with the (stacked) change in log hours worked as the dependent variable.[26] In columns (7), (8) and (9) we find that increased internet sales are associated with an absolute decline in occupational employment. We note that this is an even stronger result than suggested by the descriptive facts, which indicated slow, but positive, growth for the subset of occupations most impacted by internet sales. In other words, since employment growth was positive, on average, over the period, the effect due to increased exposure to internet sales *offset* overall employment growth. We further note that the negative effect estimated here is on the order of 10 percent of the observed, positive employment growth.

In summary, the results indicate that the rise in scale of operation due to the spread of the internet has been associated with an increase in within-occupation inequality, a decline in employment, and a fall in wages at the bottom of the distribution for the most impacted occupations, consistent with Propositions 2 and 4 and Corollary 1.

7 Concluding Remarks

A key aspect of Information and Communication Technologies (ICT) is that they increase the size of the market – or the "scale of operation" – for workers in some occupations. This paper explores the consequences due to this unique aspect of ICT for income, the income

[24] In unreported results we additionally control for initial period employment in each occupation, but this has a negligible effect on the estimates.

[25] The mean of the internet exposure measure rose from 0 to 0.05 over the period. Multiplied by the coefficient in column (3) from Table 2 gives a value of 0.014, which is 39 percent of the total rise in the 90-10 wage gap over the period. A similar calculation leads to the value with respect to the 90-50 gap.

[26] In this case we control for the pre-trend in the 90-10 wage gap, but the results are nearly identical to controlling for the pre-trend in the 90-50 gap.

distribution, and occupational choice of workers within affected occupations. We model the scale of operation as the limit up to which the production technology displays increasing returns to scale. We find that by enlarging the scale of operation in the affected occupation, ICT intensifies competition between workers and thereby lowers the price of the services they provide. It also increases the log wage gap between workers in these occupations. Lastly, despite its direct role in complementing worker output, it simultaneously drives the lowest-ability workers out of affected occupations and reduces the earnings of the next lowest-ability workers (under certain conditions) due to a competition effect. This effect highlights the unique nature of ICT innovation since this effect is not generated by other types of technological change – for instance, innovation that reduces the marginal cost of production.

We conclude by noting that there are clearly many other types of technological changes, and each may have different implications for the labor market. Furthermore, there are a range of forces, both technological and otherwise, that have contributed to the rising inequality observed in many countries in recent decades. We believe that the technological forces that we consider here are important in part due to their near ubiquity, as well as the fact that they may be relatively difficult for policy-makers to counter compared to institutional factors, such as the extent of unionization or tax policies.

References

Daron Acemoglu and David Autor. Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics*, volume 4, pages 1043–1171. Elsevier, 2011a.

Daron Acemoglu and David Autor. Skills, Tasks and Technologies: Implications for Employment and Earnings. In O. Ashenfelter and D. Card, editors, *Handbook of Labor Economics*, volume 4, chapter 12, pages 1043–1171. Elsevier, 2011b.

Sangmin Aum, Sang Yoon Tim Lee, and Yongseok Shin. Computerizing industries and routinizing jobs: Explaining trends in aggregate productivity. *Journal of Monetary Economics*, 97:1–21, 2018.

D Autor and David Dorn. The growth of low-skill service jobs and the polarization of the us labor market. *The American Economic Review*, 103(5):1553–1597, 2013.

David H Autor and Michael J Handel. Putting tasks to the test: Human capital, job tasks, and wages. *Journal of Labor Economics*, 31(S1):S59–S96, 2013.

David H. Autor, Frank Levy, and Richard J. Murnane. The Skill Content Of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, 118(4): 1279–1333, November 2003.

David H Autor, Lawrence F Katz, and Melissa S Kearney. Trends in us wage inequality: Revising the revisionists. *The Review of economics and statistics*, 90(2):300–323, 2008.

Jon Bakija, Adam Cole, and Bradley T Heim. Jobs and income growth of top earners and the causes of changing income inequality: Evidence from us tax return data. *Williamstown: Williams College*, 2010.

Richard Baldwin and James Harrigan. Zeros, Quality and Space: Trade Theory and Trade Evidence. Technical report, National Bureau of Economic Research, 2007.

Paul Beaudry, Mark Doms, and Ethan Lewis. Should the Personal Computer Be Considered a Technological Revolution? Evidence from U.S. Metropolitan Areas. *Journal of Political Economy*, 118(5):988 – 1036, 2010.

Eli Berman, John Bound, and Stephen Machin. Implications of skill-biased technological change: international evidence. *The quarterly journal of economics*, 113(4):1245–1279, 1998.

Christian Broda and David E Weinstein. Globalization and the gains from variety. *The Quarterly journal of economics*, 121(2):541–585, 2006.

David Card and John E DiNardo. Skill-biased technological change and rising wage inequality: Some problems and puzzles. *Journal of labor economics*, 20(4):733–783, 2002.

Wen-Hao Chen, Michael Forster, and Ana Llena-Nozal. Globalisation, Technological Progress and Changes in Regulations and Institutions–Which Impact on the Rise of Earnings Inequality in OECD Countries? *LIS Working Paper Series*, 2013.

Guido Matias Cortes. Where have the middle-wage workers gone? a study of polarization using panel data. *Journal of Labor Economics*, 34(1):63–105, 2016.

Arnaud Costinot and Jonathan Vogel. Matching and inequality in the world economy. *Journal of Political Economy*, 118(4):747–786, 2010.

Maya Eden and Paul Gaggl. On the welfare implications of automation. *Review of Economic Dynamics*, 29:15–43, 2018.

Hartmut Egger and Udo Kreickemeier. Fairness, Trade, and Inequality. *Journal of International Economics*, 86(2):184–196, 2012.

Sergio Firpo, Nicole M Fortin, and Thomas Lemieux. Occupational Tasks and Changes in the Wage Structure. Textos para discussão 284, Escola de Economia de São Paulo, Getulio Vargas Foundation (Brazil), September 2012.

Xavier Gabaix and Augustin Landier. Why has CEO Pay Increased So Much? Technical report, National Bureau of Economic Research, 2006.

Luis Garicano and Esteban Rossi-Hansberg. Inequality and the Organization of Knowledge. *American Economic Review*, pages 197–202, 2004.

Luis Garicano and Esteban Rossi-Hansberg. The knowledge economy at the turn of the twentieth century: the emergence of hierarchies. *Journal of the European Economic Association*, 4(2-3):396–403, 2006.

Luis Garicano and Esteban Rossi-Hansberg. Knowledge-based hierarchies: Using organizations to understand the economy. Technical report, National Bureau of Economic Research, 2014.

Charles I Jones and Jihee Kim. Exploring the dynamics of top income inequality. Technical report, mimeo.(2013)?A Schumpeterian Model of Top Income Inequality,? mimeo, 2012.

Jaewon Jung and Jean Mercenier. On modeling task, skill and technology upgrading effects of globalization with heterogeneous labor. *Economic Modelling*, 39:49–62, 2014.

Paul R Krugman. Increasing returns, monopolistic competition, and international trade. *Journal of international Economics*, 9(4):469–479, 1979.

Maurice Kugler and Eric Verhoogen. Prices, Plant Size, and Product Quality. *The Review of Economic Studies*, 79(1):307–339, 2012.

Sang Yoon Lee and Yongseok Shin. Horizontal and vertical polarization: Task-specific technological change in a multi-sector economy. 2017.

Robert E Lucas Jr. On the Size Distribution of Business Firms. *The Bell Journal of Economics*, pages 508–523, 1978.

Marc J Melitz. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6):1695–1725, 2003.

Marc J Melitz and Gianmarco IP Ottaviano. Market size, trade, and productivity. *The review of economic studies*, 75(1):295–316, 2008.

Ferdinando Monte. Skill Bias, Trade, and Wage Dispersion. *Journal of International Economics*, 83(2):202–218, 2011.

Derek Neal and Sherwin Rosen. Theories of the Distribution of Earnings. In A.B. Atkinson and F. Bourguignon, editors, *Handbook of Income Distribution*, volume 1 of *Handbook of Income Distribution*, chapter 7, pages 379–427. Elsevier, 00 2000.

Sherwin Rosen. The Economics of Superstars. *American Economic Review*, 71(5):845–58, December 1981.

Sherwin Rosen. Specialization and Human Capital. *Journal of Labor Economics*, 1(1):43–49, January 1983.

Steven J. Ruggles, Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Mathew Sobek. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Technical report, Minneapolis: University of Minnesota, 2010.

Gilles Saint-Paul. Distribution and Growth in an Economy with Limited Needs: Variable Markups and the End of Work. *The Economic Journal*, 116(511):382–407, 2006.

Gilles Saint-Paul. Knowledge Hierarchies in the Labor Market. *Journal of Economic Theory*, 137(1):104–126, 2007.

Michael Sattinger. Assignment Models of the Distribution of Earnings. *Journal of Economic Literature*, pages 831–880, 1993.

Anson Soderbery. Estimating import supply and demand elasticities: Analysis and implications. *Journal of International Economics*, 96(1):1–17, 2015.

Jan Tinbergen. Substitution of graduate by other labour*. *Kyklos*, 27(2):217–226, 1974.

	Overall			Occs Most Exposed to Internet Sales vs. Other Occs			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Year	Total Wage Dispersion	Within-Occupation	Between-Occupation	Within (Top 10% Most Exposed)	Within (Other)	Between (Top 10% Most Exposed)	Between (Other)
1990	0.298	0.223	0.075	0.028	0.195	0.013	0.062
2000	0.322	0.233	0.089	0.032	0.200	0.014	0.075
2010	0.377	0.254	0.123	0.048	0.206	0.022	0.101

Table 3: Wage Dispersion Within and Between Occupations, 1990-2010

APPENDIX

A Decomposition of Wage Dispersion

We begin by decomposing aggregate log wage dispersion into within- and between-occupation components separately for 1990, 2000 and 2010. Formally, we calculate:

$$\frac{1}{N_t}\sum_i (w_{it}-\bar{w}_t)^2 = \frac{1}{N_t}\sum_l \sum_{i\in l}(w_{it}-\bar{w}_{lt})^2 + \frac{1}{N_t}\sum_l N_{lt}(\bar{w}_{lt}-\bar{w}_t)^2 \tag{23}$$

where workers are indexed by i and the year by t; l represents occupations; N_{lt} and N_t represent the number of workers in each occupation and overall; and w_{it}, $\bar{w}_{lt}$ and $\bar{w}_t$ are the log worker wage, the average log occupational wage, and the overall average wage. In using the log wage we ensure the values are independent of the wage units. The first term on the right hand side reflects the within-occupation component of wage inequality.

Table 3 reports the results. The first three columns of the table highlight the overall contributions of within- and between-occupation inequality. First, throughout the period the contribution of within-occupation inequality to aggregate inequality in any particular year is large relative to the between component. Furthermore, 40 percent of the rise in aggregate inequality between 1990 and 2010 was due to a rise in within-occupation inequality (Fact 1). We can decompose total log wage dispersion further by noting that the within term in (23) is the sum across individual occupations, and so the contribution of different subsets of occupations can be easily separated out. As it turns out, most of the rise in within-occupation inequality was due to a particular subset of occupations, namely those most affected by the internet. Note that we define internet exposure according to (21).

The last four columns of Table 3 once again decompose total log wage dispersion in each year into the within and between components, but then decompose each of these into two further sets of occupations reflecting 1) the top 10 percent of occupations according to measure (21) and 2) all other occupations. Comparing columns (2) and (4), we see that 65 percent of the rise in within-occupation wage inequality between 1990 and 2010 is due to the set of occupations that were most exposed to internet sales (Fact 2).

B Proofs

B.1 Proof of Proposition 1

Proof. We first establish that aggregate spending on this good is $BH_k(\mu P)^{\frac{1}{1-\widehat{\rho}}}$. To see this, note that from (3) it follows that the aggregate income spent on the alternative good is $M \times \left[1+\mu^{1/(\widehat{\rho}-1)}P^{\widehat{\rho}/(\widehat{\rho}-1)}\right]^{-1}$. To find the aggregate income M, observe that with the price of each variety given by (13), the price index, from (11), is

$$P = \left(\frac{Mf(P,\mu)}{B}\right)^{1-\rho} H_k^{\rho-1}. \tag{24}$$

With $f(P,\mu) = P^{\frac{\rho-\widehat{\rho}}{(1-\rho)(1-\widehat{\rho})}} / \left(\mu^{\frac{1}{1-\widehat{\rho}}} + P^{\frac{\widehat{\rho}}{\widehat{\rho}-1}}\right)$, it follows that

$$M = BPH_k \times \left[1+\mu^{\frac{1}{1-\widehat{\rho}}}P^{\frac{\widehat{\rho}}{1-\widehat{\rho}}}\right]. \tag{25}$$

Therefore, aggregate spending on the alternative good is

$$BPH_k \times \left[1+\mu^{\frac{1}{1-\widehat{\rho}}}P^{\frac{\widehat{\rho}}{1-\widehat{\rho}}}\right] \times \left[1+\mu^{1/(\widehat{\rho}-1)}P^{\widehat{\rho}/(\widehat{\rho}-1)}\right]^{-1} = BH_k(\mu P)^{\frac{1}{1-\widehat{\rho}}}$$

Next, we find the aggregate supply of the alternative good. A mass $1-k$ of agents provide services, each demanding Bc/A units of labor as input. The total labor supply is k. Thus, $k - \frac{c}{A}B \times (1-k)$ agents work to produce the alternative good, yielding an output of $[k - \frac{c}{A}B \times (1-k)] \times A = kA - (1-k)cB$. Note that this aggregate supply of the alternative good can be re-written as $(A+Bc)(k-k_0)$, with $k_0 = \frac{Bc}{A+Bc}$ is the threshold for the number of agents at which the aggregate supply of the alternative good is zero.

Market clearing for the alternative good thus implies:

$$BH_k(\mu P)^{\frac{1}{1-\widehat{\rho}}} = (A+Bc)(k-k_0). \tag{26}$$

This equation contains P, an endogenous variable. To determine k, we exploit an additional connection between P and k, as follows. Using (24) to cancel $\left(\frac{Mf(P,\mu)}{B}\right)^{1-\rho}$ in (13), we find $p_i = PH_k^{1-\rho}h_i^{\rho}$ for any $i \geq k$, in particular when $i = k$. At the same time, in (15) we found $p_k = A/B + c$. Therefore,

$$PH_k^{1-\rho}h_k^{\rho} = \frac{A}{B} + c. \tag{27}$$

Solving for P, substituting it into (26) and rearranging, we arrive at equation (18) that pins down k in equilibrium.

The right hand side term – the aggregate supply of the alternative good – increases with k from 0 to $1 - k_0 > 0$ over $k \in [k_0, 1]$. At the same time, the left hand side term – the aggregate spending on the alternative good – (1) decreases with k and (2) reaches 0 at $k = 1$, hence a unique root for k over $[k_0, 1]$. To see (1), observe that since $\rho - \widehat{\rho} > 0$, $H_k^{\frac{\rho-\widehat{\rho}}{1-\widehat{\rho}}}$ increases with $H_k = \{\int_k^1 h_i^{\rho}\}^{\frac{1}{\rho}}$, which decreases with k; and that with $\rho > 0$, $h_k^{\frac{-\rho}{1-\widehat{\rho}}}$ decreases with h_k which, by assumption, increases with k. Point (2) is simply because $H_1 = 0$. ∎

B.2 Proof of Proposition 2

Proof. We intend to prove that $\frac{d\log m_i}{dB}$ increases with m_i. For a professional $i \geq k$, let $\widetilde{m}_i := m_i + Bc$ be his revenue. Then $\frac{d\log m_i}{dB} = \frac{1}{m_i} \times \frac{dm_i}{dB} = \frac{1}{m_i} \times \left[\frac{d\widetilde{m}_i}{dB} - c\right] = \frac{1}{m_i} \times \left[\frac{d\log\widetilde{m}_i}{dB}\widetilde{m}_i - c\right]$. Therefore,

$$\frac{d\log m_i}{dB} = \frac{d\log\widetilde{m}_i}{dB} \times \left(1 + \frac{Bc}{m_i}\right) - \frac{c}{m_i}. \tag{28}$$

From (16),

$$\widetilde{m}_i = (Bc + A)\frac{h_i^{\rho}}{h_k^{\rho}}.$$

Thus $\log\widetilde{m}_i = \log(Bc + A) + \rho\log h_i - \rho\log h_k$. It follows that

$$\frac{d\log\widetilde{m}_i}{dB} = \frac{c}{Bc + A} - \rho\left[\log h_k\right]_k' \frac{dk}{dB} \tag{29}$$

and is independent of m_i. Hence, from (28), $\left[\frac{d\log m_i}{dB}\right]_{m_i}' = \frac{d\log\widetilde{m}_i}{dB} \times \left[1 + \frac{Bc}{m_i}\right]_{m_i}' - \left[\frac{c}{m_i}\right]_{m_i}' = \frac{d\log\widetilde{m}_i}{dB} \times \frac{-Bc}{m_i^2} + \frac{c}{m_i^2}$. It follows that $\left[\frac{d\log m_i}{dB}\right]_{m_i}' > 0$ if and only if:

$$\frac{d\log\widetilde{m}_i}{dB} < \frac{1}{B}. \tag{30}$$

Using equation (29), this equality is equivalent to $\left[\frac{c}{Bc+A} - \rho\,[\log h_k]'_k \frac{dk}{dB}\right] B < 1 \Leftrightarrow$

$$B\rho\,[\log h_k]'_k \left(-\frac{dk}{dB}\right) < \frac{A}{Bc+A}. \tag{31}$$

Because $k_0 = \frac{Bc}{A+Bc}$ and hence $\frac{dk_0}{dB} = \frac{Ac}{(A+Bc)^2}$, we have $\frac{dk}{dB} = \frac{dk}{dk_0}\frac{Ac}{(A+Bc)^2}$. Substitute this into inequality (31), and the inequality becomes:

$$\rho\,[\log h_k]'_k\, k_0 \left(-\frac{dk}{dk_0}\right) < 1. \tag{32}$$

The inequality is true as long as $\frac{dk}{dk_0} > 0$. Next we show that it is also true if $\frac{dk}{dk_0} < 0$, using equilibrium condition (18) to find an upper bound for $-\frac{dk}{dk_0}$. Because $A/B + c = c/k_0$, this condition can be rearranged as

$$\mu^{\frac{1}{1-\hat{\rho}}} H_k^{\frac{\rho-\hat{\rho}}{1-\hat{\rho}}} h_k^{\frac{-\rho}{1-\hat{\rho}}} = (k-k_0)\,(\frac{k_0}{c})^{\frac{\hat{\rho}}{1-\hat{\rho}}}. \tag{33}$$

Taking the logarithm on both sides we have:

$$-\frac{dk}{dk_0} = \frac{\frac{\hat{\rho}k-k_0}{(1-\hat{\rho})k_0(k-k_0)}}{\frac{1}{k-k_0} + \frac{\rho}{1-\hat{\rho}}\,[\log h_k]'_k - \frac{\rho-\hat{\rho}}{1-\hat{\rho}}\,[\log H_k]'_k}. \tag{34}$$

By assumption, $[\log h_k]'_k > 0$. Since (14) implies that $H_x^\rho = \int_x^1 h_i^\rho$ decreases with x, we have $[\log H_k]'_k < 0$. Furthermore, by assumption $\frac{\rho-\hat{\rho}}{1-\hat{\rho}} > 0$. It follows that all three terms in the denominator of (34) are positive. $-\frac{dk}{dk_0} > 0$ if and only if the numerator is positive. If it is positive then we have

$$-\frac{dk}{dk_0} < \frac{\hat{\rho}k - k_0}{k_0\,(k-k_0) \times \rho\,[\log h_k]'_k}.$$

As a result, if $-\frac{dk}{dk_0} > 0$, inequality (32) follows from $\rho\,[\log h_k]'_k\, k_0 \times \frac{\hat{\rho}k-k_0}{k_0(k-k_0)\times\rho[\log h_k]'_k} \le 1$, which is equivalent to $\hat{\rho}k - k_0 \le k - k_0$, which indeed is true. ∎

B.3 Proof of Proposition 3

Proof. Because $m_i = (p_i - c)\,B$. Therefore, $\frac{d\log m_i}{dB} = \frac{d\log(p_i-c)}{dB} + \frac{1}{B}$. Hence, the proposition is equivalent to

$$\frac{d\log m_i}{dB} < \frac{1}{B}. \tag{35}$$

In the proof of Proposition 3, we have proved that $\frac{d\log\widetilde{m}_i}{dB} < \frac{1}{B}$ in (30). To prove Proposition 3, it therefore suffices to prove that

$$\frac{d\log m_i}{dB} < \frac{d\log\widetilde{m}_i}{dB},$$

which, by (28), is equivalent to $\frac{d\log\widetilde{m}_i}{dB} \times \frac{Bc}{m_i} - \frac{c}{m_i} < 0$, or to $\frac{d\log\widetilde{m}_i}{dB} < \frac{1}{B}$, which has been proved. ∎

B.4 Proof of Proposition 4

Proof. Because $\frac{dk}{dB} = \frac{dk}{dk_0}\frac{Ac}{(A+Bc)^2}$, we have $\frac{dk}{dB} > 0$ if and only if $\frac{dk}{dk_0} > 0$, which, by equation (34) (see the proof of Proposition 3), holds true if and only if

$$\widehat{\rho}k < k_0. \tag{36}$$

This inequality certainly holds if $\widehat{\rho} \leq k_0$ because $k < 1$. We now prove that it also holds true if $\widehat{\rho} > k_0$ and μ is above a threshold. The equilibrium cutoff k is determined by equation (33) (see the proof of Proposition 3), which is equivalent to

$$\mu = (\frac{k_0}{c})^{\widehat{\rho}} (k - k_0)^{1-\widehat{\rho}} \, h_k^{\rho} H_k^{-(\rho-\widehat{\rho})} := g(k).$$

Observe that $g' > 0$ because $1-\widehat{\rho} > 0$, $\rho > 0$, $h_k' > 0$, $\rho-\widehat{\rho} > 0$ and H_k decreases with k. Hence, $k < k_0/\widehat{\rho}$ – namely, inequality (36) holds – if and only if $g(k) < g(k_0/\widehat{\rho})$, that is, $\mu < g(k_0/\widehat{\rho})$. Also, observe that $g(k_0/\widehat{\rho}) = \eta\left(\frac{h_{k_0/\widehat{\rho}}}{H_{k_0/\widehat{\rho}}}\right)^{\rho}$, where $\eta := (\frac{k_0}{c})^{\widehat{\rho}}(k_0/\widehat{\rho} - k_0)^{1-\widehat{\rho}} H_{k_0/\widehat{\rho}}^{\widehat{\rho}}$ is independent of ρ and $\eta = 0$ if $k_0 = 0$. Hence (34) holds – that is $\frac{dk}{dB} > 0$ – if and only if either $\widehat{\rho} \leq k_0$ or $k_0 > 0$ and $\mu < \eta\left(\frac{h_{k_0/\widehat{\rho}}}{H_{k_0/\widehat{\rho}}}\right)^{\rho}$. ∎

B.5 Proof of Corollary 1

Proof. From the discussion between equations (14) and (15), we know that $p_i = p_k \times h_i^{\rho}/h_k^{\rho}$ and $p_k = A/B+c$. Thus $d\log p_i/dB = d\log p_k/dB - d\log h_k^{\rho}/dB < -d\log h_k^{\rho}/dB = -(\rho h_k'/h_k \times dk/dB) < 0$, as $h_k' > 0$ and $dk/dB > 0$ by Proposition 3. ∎

B.6 Proof of Proposition 5

The proof follows from the two lemmas and their proofs below.

Lemma 11. *Assume that $\Omega(\rho) \cdot A/(A+Bc) < 1$ and $\widehat{\rho} \geq 0$. Then $dm_i/dB > 0$, namely agent i's income rises with an increase in the limit of IRS as long as*

$$\frac{h_i}{h_k} > \left(\frac{1}{1-\Omega(\rho) \cdot A/(A+Bc)}\right)^{\frac{1}{\rho}}. \tag{37}$$

Proof of Lemma 1

Proof. By (16),

$$\frac{dm_i}{dB} = h_i^\rho/h_k^\rho \cdot [c-(A+Bc) \cdot \rho \cdot (\log h_k)' \cdot \frac{dk}{dB}] - c. \tag{38}$$

The identity of the marginal professional, k, is determined by equation (46). Taking the logarithm of both sides: $\frac{1}{1-\widehat{\rho}} \log \mu + \frac{\rho-\widehat{\rho}}{\rho(1-\widehat{\rho})} \log H_k^\rho - \frac{\rho}{1-\widehat{\rho}} \log h_k = \log(k-k_0) - \frac{\widehat{\rho}}{1-\widehat{\rho}} \log(A/B+c)$. Now taking the derivative with respect to B on both sides and noting that $\frac{dH_k^\rho}{dk} = -h_k^\rho$ and recalling $k_0 = \frac{Bc}{A+Bc}$: $[-\frac{\rho-\widehat{\rho}}{\rho(1-\widehat{\rho})} h_k^\rho/H_k^\rho - \frac{\rho}{1-\widehat{\rho}}(\log h_k)'] \cdot \frac{dk}{dB} = \frac{1}{k-k_0} \cdot [\frac{dk}{dB} - \frac{Ac}{(A+Bc)^2}] + \frac{\widehat{\rho}}{1-\widehat{\rho}} \cdot \frac{A}{(A+Bc)B} \Rightarrow$

$$\frac{dk}{dB} = \frac{1/(k-k_0) \cdot Ac/(A+Bc)^2 - \widehat{\rho}/(1-\widehat{\rho}) \cdot A/[A+Bc)B]}{1/(k-k_0) + \frac{\rho}{1-\widehat{\rho}}(\log h_k)' + \frac{\rho-\widehat{\rho}}{\rho(1-\widehat{\rho})} h_k^\rho/H_k^\rho}.$$

The numerator is smaller than $1/(k-k_0) \cdot Ac/(A+Bc)^2$, while the denominator is greater than $1/(k-k_0) + \frac{\rho}{1-\widehat{\rho}}(\log h_k)'$, which is in turn greater than $1/(k-k_0) + \rho(\log h_k)'$. Therefore,

$$\frac{dk}{dB} < \frac{Ac/(A+Bc)^2}{1+\rho(\log h_k)'(k-k_0)}. \tag{39}$$

By (38), $\frac{dm_i}{dB} > 0$ if

$$h_i^\rho/h_k^\rho \cdot [c-(A+Bc) \cdot \rho \cdot (\log h_k)' \cdot \frac{dk}{dB}] > c. \tag{40}$$

With an upper bound of $\frac{dk}{dB}$ given by (39), this inequality follows from: $h_i^\rho/h_k^\rho \cdot [c-(A+Bc) \cdot \rho \cdot (\log h_k)' \cdot \frac{Ac/(A+Bc)^2}{1+\rho(\log h_k)'(k-k_0)}] > c \Leftrightarrow$

$$h_i^\rho/h_k^\rho \cdot [1 - \frac{A}{A+Bc} \cdot \frac{\rho \cdot (\log h_k)'}{1+\rho(\log h_k)'(k-k_0)}] > 1, \tag{41}$$

which is equivalent to (37). ∎

Condition (37), however, is not easy to check. This is because k is determined in equilibrium and depends on the distribution of human capital (specifically, the functional form of $h(i)$). We therefore present an approach, dispensing with k, to obtain a condition under which the top professionals gain on net from an increase in the limit of IRS.

Lemma 12. *Assume* $h_1 > 1$. *If for some* ζ, $h_1 \geq \zeta \cdot h(f(k_0, D \cdot \zeta^{\frac{\rho}{1-\widehat{\rho}}}))$, *then* $h_1 > \zeta \cdot h_k$.

Proof of Lemma 2

Proof. We prove the lemma in three steps.

Step 1: If $h_1 > 1$, then

$$k - k_0 < D(\frac{h_1}{h_k})^{\frac{\rho}{1-\widehat{\rho}}}(1-k)^{\frac{\rho-\widehat{\rho}}{\rho(1-\widehat{\rho})}}. \tag{42}$$

Proof: k is determined by equation (46), or equivalently, $k - k_0 = DH_k^{\frac{\rho-\widehat{\rho}}{1-\widehat{\rho}}} h_k^{\frac{-\rho}{1-\widehat{\rho}}}$. Note that $H_k = \{\int_k^1 h_i^\rho\}^{\frac{1}{\rho}}|_{h_i'>0} < \{\int_k^1 h_1^\rho\}^{\frac{1}{\rho}} = h_1(1-k)^{\frac{1}{\rho}}$. Therefore, $H_k^{\frac{\rho-\widehat{\rho}}{1-\widehat{\rho}}} h_k^{\frac{-\rho}{1-\widehat{\rho}}} = (\frac{H_k^{\rho-\widehat{\rho}}}{h_k^\rho})^{\frac{1}{1-\widehat{\rho}}} < (\frac{h_1^{\rho-\widehat{\rho}}(1-k)^{\frac{\rho-\widehat{\rho}}{\rho}}}{h_k^\rho})^{\frac{1}{1-\widehat{\rho}}} = h_1^{\frac{\rho-\widehat{\rho}}{1-\widehat{\rho}}}/h_k^{\frac{\rho}{1-\widehat{\rho}}} \cdot (1-k)^{\frac{\rho-\widehat{\rho}}{\rho(1-\widehat{\rho})}}|_{\frac{\rho-\widehat{\rho}}{1-\widehat{\rho}} \leq \frac{\rho}{1-\widehat{\rho}} \text{ and } h_1>1} < h_1^{\frac{\rho}{1-\widehat{\rho}}}/h_k^{\frac{\rho}{1-\widehat{\rho}}} \cdot (1-k)^{\frac{\rho-\widehat{\rho}}{\rho(1-\widehat{\rho})}}$, which implies (42).

Step 2:

$$k < f(k_0, D \cdot (\frac{h_1}{h_k})^{\frac{\rho}{1-\widehat{\rho}}}). \tag{43}$$

Proof: Let $\tau := f(k_0, D \cdot (\frac{h_1}{h_k})^{\frac{\rho}{1-\widehat{\rho}}})$. By the definition of $f(\cdot,\cdot)$, $\tau - k_0 = D \cdot (\frac{h_1}{h_k})^{\frac{\rho}{1-\widehat{\rho}}} \cdot (1-\tau)^{\frac{\rho-\widehat{\rho}}{\rho(1-\widehat{\rho})}}$. The two sides of this inequality minus, respectively, the two sides of inequality (42) leads to $\tau - k > D(\frac{h_1}{h_k})^{\frac{\rho}{1-\widehat{\rho}}}[(1-\tau)^{\frac{\rho-\widehat{\rho}}{\rho(1-\widehat{\rho})}} - (1-k)^{\frac{\rho-\widehat{\rho}}{\rho(1-\widehat{\rho})}}]$. This inequality can hold true only if $\tau > k$: if $\tau \leq k$, then the LHS of the inequality is negative, while the RHS is positive – and thus cannot be strictly smaller than the LHS – because $1-\tau \geq 1-k$, which implies $(1-\tau)^{\frac{\rho-\widehat{\rho}}{\rho(1-\widehat{\rho})}} - (1-k)^{\frac{\rho-\widehat{\rho}}{\rho(1-\widehat{\rho})}} \geq 0$ (as $\frac{\rho-\widehat{\rho}}{\rho(1-\widehat{\rho})} > 0$). Q.E.D.

Step 3: We prove the Lemma by showing that $\zeta \geq h_1/h_k$ leads to a contradiction. Clearly, $f(k_0, y)$ increases with y, and therefore if $\zeta \geq h_1/h_k$, then $f(k_0, D \cdot (\frac{h_1}{h_k})^{\frac{\rho}{1-\widehat{\rho}}}) < (f(k_0, D \cdot \zeta^{\frac{\rho}{1-\widehat{\rho}}})$, which together with (43) implies that $k < f(k_0, D \cdot \zeta^{\frac{\rho}{1-\widehat{\rho}}})$. Since $h'(i) > 0$, then $h_k < h(f(k_0, D \cdot \zeta^{\frac{\rho}{1-\widehat{\rho}}}))$. Thus we have

$$\zeta \geq \frac{h_1}{h_k} > \frac{h_1}{h(f(k_0, D \cdot \zeta^{\frac{\rho}{1-\widehat{\rho}}}))},$$

which implies $\zeta \cdot h(f(k_0, D \cdot \zeta^{\frac{\rho}{1-\widehat{\rho}}})) > h_1$, in contradiction to the lemma. ∎

B.7 Proof of Proposition 6

Proof. For professionals $i < i^*$, their capacity constraint is non-binding and the optimal price for them is thus $p_i = c/\rho$. From equations (6) and (8), $m_i = \varpi(B) \times h_i^{\rho/(1-\rho)}$ for some function $\varpi(B)$. Hence, $\frac{d\log m_i}{dB} = \frac{d\log \varpi(B)}{dB}$ and is independent of m_i or weakly increasing with it. For professionals $i > i^*$, their capacity constraint is binding. Thus equation (13) holds. Following the discussion ensuing this equation we find that $p_i = p_{i^*} \times h_i^\rho / h_{i^*}^\rho$ and hence $m_i = (p_c - c) B = p_{i^*} B \times h_i^\rho / h_{i^*}^\rho - Bc$. As the capacity constraint just starts binding at i; we know that $p_{i^*} = c/\rho$ and independent of B. Following the proof of Proposition 3, we let $\widetilde{m}_i := m_i + Bc = p_{i^*} B \times h_i^\rho / h_{i^*}^\rho$ be the revenue. Then

$$\frac{d\log \widetilde{m}_i}{dB} = \frac{1}{B} - \rho \frac{h'_{i^*}}{h_{i^*}} \frac{di^*}{dB}. \tag{44}$$

The proof of Proposition 3 has shown that $\left[\frac{d\log m_i}{dB}\right]'_{m_i} > 0$ if and only if:

$$\frac{d\log \widetilde{m}_i}{dB} < \frac{1}{B}, \tag{45}$$

which, with (44), holds true because $\frac{di^*}{dB} > 0$, that is, an increase in B make the capacity constraint becomes non-binding for more professionals. ∎

B.8 Proof of Proposition 7

Proof. For $i \leq i^*$, $p_i = c/\rho$ and thus $\frac{dp_i}{dB} = 0$. For $i > i^*$, following the proof of Proposition 4, $\frac{dp_i}{dB} < 0$ if $\frac{d\log \widetilde{m}_i}{dB} < \frac{1}{B}$, which is proved as inequality (45) in the proof of Proposition 8. ∎

B.9 Proof of Proposition 8

Proof. Equilibrium condition (18) is equivalent to

$$\mu^{\frac{1}{1-\widehat{\rho}}} H_k^{\frac{\rho-\widehat{\rho}}{1-\widehat{\rho}}} h_k^{\frac{-\rho}{1-\widehat{\rho}}} = (A/B + c)^{\frac{-\widehat{\rho}}{1-\widehat{\rho}}} \left[k - Bc/(A + Bc)\right]. \tag{46}$$

Use the implict function theorem and recall that $k_0 = Bc/(A + Bc)$, and we have:

$$\frac{dk}{dc} = \frac{\frac{\widehat{\rho}}{1-\widehat{\rho}} \frac{B}{A+Bc} + \frac{1}{k-k_0} \frac{AB}{(A+Bc)^2}}{\frac{1}{k-k_0} + \frac{\rho}{1-\widehat{\rho}} \left[\log h_k\right]'_k - \frac{\rho-\widehat{\rho}}{1-\widehat{\rho}} \left[\log H_k\right]'_k}.$$

The denominator, as we argued in the proof of Proposition 5, is positive. Hence, $\frac{dk}{dc} > 0$ if and only if the numerator is positive, which is equivalent to

$$\frac{\widehat{\rho}}{1-\widehat{\rho}} + \frac{1-k_0}{k-k_0} > 0,$$

which holds true because $\frac{\widehat{\rho}}{1-\widehat{\rho}} + \frac{1-k_0}{k-k_0} > \frac{\widehat{\rho}}{1-\widehat{\rho}} + 1 = \frac{1}{1-\widehat{\rho}} > 0$.

For the second claim of the proposition, we follow the approach used to prove Proposition 5 and show that $\left[\frac{d\log m_i}{dc}\right]'_{m_i} > 0$. Let $\widetilde{m}_i := m_i + Bc$. Then

$$\frac{d\log m_i}{dc} = \frac{d\log \widetilde{m}_i}{dc} \times \frac{\widetilde{m}_i}{m_i} - \frac{B}{m_i}. \tag{47}$$

From (16),

$$\widetilde{m}_i = (Bc + A)\frac{h_i^\rho}{h_k^\rho}.$$

Thus $\log \widetilde{m}_i = \log(Bc + A) + \rho \log h_i - \rho \log h_k$ and

$$\frac{d\log \widetilde{m}_i}{dc} = \frac{B}{Bc + A} - \rho\frac{h'_k}{h_k}\frac{dk}{dc}. \tag{48}$$

Observe that $\frac{d\log \widetilde{m}_i}{dc}$ is independent of m_i. Hence, from (47), $\left[\frac{d\log m_i}{dB}\right]'_{m_i} = \frac{d\log \widetilde{m}_i}{dc} \times \left[\frac{m_i+Bc}{m_i}\right]'_{m_i} - \left[\frac{B}{m_i}\right]'_{m_i} = \frac{d\log \widetilde{m}_i}{dc} \times \frac{-Bc}{m_i^2} + \frac{B}{m_i^2}$. It follows that $\left[\frac{d\log m_i}{dc}\right]'_{m_i} > 0$ if and only if $-\frac{d\log \widetilde{m}_i}{dc} \times c + 1 > 0$, which, with (29), is equivalent to $-\frac{Bc}{Bc+A} + \rho c \frac{h'_k}{h_k}\frac{dk}{dc} + 1 > 0$, which certainly holds true as $\frac{dk}{dc} > 0$. ∎