Increasing Returns to Scale Within Limits: A Model of ICT

Tianxi Wang and Greg C. Wright*

February 2019

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 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 often 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, ICT is also special: by increasing the scale of operation for *all* workers it may intensify competition between them, leading to a countervailing indirect effect. 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 is based on the observation that for occupations in which scale of operation is an important determinant of workers' income, the production technology displays Increasing Returns to Scale up to some Limit (IRSL).[3] For example, while it is costly to produce a song with mass appeal, it costs little to admit an additional person into the theater to hear it, up to the point that the theater is filled. This feature of large fixed costs and small marginal costs, over some range, is reflective of the presence of IRS up to a limit, which is the theater's capacity. Similarly, the capacity of the stadium defines the limit of IRS for football players prior to the invention of television while booksellers faced a limit defined by their local market prior to the internet. We therefore model the scale of operation

[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, and would not be possible otherwise.

[3] Taking the limit to infinity, IRSL subsumes IRS as a special case.

as the limit up to which the production technology displays IRS.

In our model, agents choose to work in either a *professional* occupation or an *alternative* occupation. The former represents all occupations in which the scale of operation is increased by ICT, while the latter represents all other occupations. The professional occupation therefore exhibits IRSL. Formally, taking the opportunity cost of their career choice as a fixed cost, each professional produces a unique variety of services at a constant marginal cost up to some technological 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. To allow for flexible competition between them, we assume that the market for professionals' services is monopolistically competitive. 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.

Our main findings are as follows. First, an increase in the scale of operation, B, reduces the price of all professional services by intensifying competition between them. This is the indirect effect discussed above – and it is special in that it differs from the typical view in which technological change directly complements or substitutes for workplace tasks.

Second, an increase in the scale of operation leads to increased inequality within the professional occupation. This is because an increase in B allows each professional to sell more output which, on the one hand, benefits them but, on the other hand, leads to fiercer competition between them. And whereas the increase in competition is felt equally among workers, 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.

Lastly, an increase in B drives the lowest-human-capital professionals out of the occupation and reduces the earnings of the next lowest as long as the substitutability between professional services and the output produced by the alternative occupation is below some threshold.[4] In short, for these low-human-capital workers the negative effect due to stronger competition dominates the positive effect due to increased market size. Interestingly, this result indicates that the substitutability between professional services and other products in the economy is more important than the substitutability between individual professional services. The intuition for this derives from the fact that the rise in B leads to an increase in the supply of professional services, which leads to a fall in their prices. These prices fall by a larger amount when the elasticity of demand for professional services with respect to their price is lower. And this elasticity is lower when professional services are less substitutable with the output produced by the alternative occupation.

To highlight the special nature of ICT we compare these findings to the effect of technological

[4] Anecdotally, with the advent of TV fewer comedians were able to earn a living by performing for local audiences, and with the advent of the internet travel agents have seen their share of U.S. employment halved over the last two decades (Bureau of Labor Statistics).

changes that reduce 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. In short, 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 structure of 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. Section 6 presents empirical evidence consistent with our results. 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 approaches the topic from a variety of perspectives.[5] The dominant theoretical approach in this literature is the theory of skill biased technological change (SBTC hereafter). 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 (2011a) have extended the standard SBTC framework to endogenize the matching of skills to tasks. Within this literature, within-occupation inequality has received little theoretical attention. Cortes (2016) presents a theoretical (and empirical) analysis of worker sorting across occupations in response to technological change, with implications for within-occupation inequality. Furthermore, Acemoglu and Autor (2011b) (Section 3.4) outline an extension of the SBTC model that can be used to consider within-occupation inequality.[6]

Whereas the SBTC literature focuses on the extent to which technological change directly com-

[5] Other theoretical studies on the effects of technology (not necessarily ICT) include Jones and Kim (2012) who 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. Related empirical work includes Autor et al. (2017) who focus on the role of "superstar" firms' responses to technological changes and globalization; and Bessen (2017) who links Information Technology investments to rising industry concentration.

[6] In general, the theoretical literature on the effect of technological change on within-occupation inequality is sparse. Hornstein et al. (2011) present a model in which identical workers are paid different wages due to labor market frictions, a framework that can be interpreted as addressing within-occupation inequality, but as their framework has no role for heterogeneity in worker ability, their channel is very different to ours. There are also contributions within the sociology literature – e.g., see Mouw and Kalleberg (2010) or Kim and Sakamoto (2008), who document the contributions of within- and between-occupation inequality to growth in overall inequality (we do something similar in Section 6).

plements or substitutes for workers in performing workplace tasks, our model highlights an additional channel in which ICT affects workers' incomes by increasing their scale of operation. Technologically, a rise in scale of operation complements these workers; and yet some of them may still lose due to the more intense competition that it induces. Our model also highlights a role for occupations as an important unit linking technological change to rising inequality. This is conceptually distinct from the SBTC frameworks in which technological change exerts an impact on factors of production (or specific workplace tasks) regardless of the occupation in which they are employed. As a result, we focus on the implications for within-occupation changes.

This paper models ICT as an increase in the market size associated with certain occupations and analyzes its effect on within-occupation inequality and workers' occupational choices. The effect of ICT, in particular the internet, is also studied in the literature on consumer search.[7] This literature models ICT as a reduction in search costs and considers the implications for prices, market shares and product variety. While falling search costs may be linked to a rise in the scale of operation, they are not equivalent. For example, the reduction in search costs due to the internet is clearly a source of increased market size, however ICT innovations such as the radio and television increased market size without reducing search costs. In this sense our approach complements this literature while covering a broader range of ICT advances. In particular, the literature typically predicts that a reduction in search costs intensifies competition between sellers, which often leads sellers to reduce prices,[8] much like an increase in market size does in our model. Moreover, some studies such as Cachon et al. (2008) find that a lower search cost induces a market expansion effect: if consumers search more then firms have access to more consumers, which induces them to offer a larger range of products. A similar effect drives them to offer more niche products, as found by Bar-Isaac et al. (2012) and Larson (2013). Besides the difference in topic and way of modeling ICT, our paper differs from this literature in that the results are driven by the balance between increased market size and increased competition, whereas this trade-off is absent in that literature.

Aside from the consumer search literature, Garicano and Rossi-Hansberg (2014), building on Lucas Jr (1978), consider the implications of ICT innovation 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. Different from our paper, in their study ICT innovation 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, channels that are distinct from those considered in our model.

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,

[7]See among others Stahl (1989), Anderson and Renault (1999), Armstrong et al. (2009), Chen and He (2011), Zhou (2014), Koulayev (2014), Larson (2013), Jung and Mercenier (2014), and Bar-Isaac et al. (2012).

[8]We note that some studies, such as Bar-Isaac et al. (2012) and Cachon et al. (2008), find a non-monotonic relationship between price and the search cost.

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),[9] though the two papers study very different issues.[10] 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 re-allocation 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).

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.[11] 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 income takes a variety of forms, such as wages,

[9] 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).

[10] 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.

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

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.

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 to subsist on her labor endowment, namely if she chooses the alternative occupation, her income is A.

The production of professional services requires both human capital and labor. Labor is to serve some auxiliary functions, e.g. as a janitor of the 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.[12] 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 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 $c > 0$ is a constant. Thus, if an agent decides to use her time to supply human capital rather than labor, thereby providing a variety of professional services, she subsequently hires labor to produce output at constant returns to scale up to the limit B. According to this production function a unit of output requires c/A units of labor input. 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.[13] The cost function associated with producing professional services is

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

[13] In the model we abstract from phyical 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.

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

Due to the existence of fixed costs F, the average cost decreases with output y until $y > B$. 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 IRS only up to the limit B. The primary modeling innovation presented here is the introduction of this limit B of IRS. We use this B to represent the maximum scale of operation for a professional – for example, the capacity of the theater in which a musician performs.

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[14]

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

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. 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.

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

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

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

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

[14] Throughout the paper the notation "di" is omitted to simplify notation.

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

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

For an 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. Out of this expenditure the portion spent on variety i is $\left(h_i/p_i\right)^{\rho/(1-\rho)} / \int_E \left(h_j/p_j\right)^{\rho/(1-\rho)} = P^{\frac{\rho}{1-\rho}} \left(h_i/p_i\right)^{\rho/(1-\rho)}$. 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

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

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 supply labor and produce the alternative good, she gets A. If the agent chooses to live on her human capital 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:

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

The agent chooses to provide professional services instead of supplying labor only if

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

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

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

If $i < k$, agent i earns income A, and if $i \geq k$ agent i earns $m(h_i)$. 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

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

and aggregate income is

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

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 supply labor if and only if $i < k$ where k is determined by (10);

(iii) Aggregate income is given by (12).[15]

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}.[16] It follows that for any $i \geq k$, $p_i/p_k = h_i^{\rho}/h_k^{\rho}$. Hence, $p_iB/\int_k^1 (p_jB) = h_i^{\rho}/\int_k^1 h_j^{\rho}$,

[15] 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.

[16] That is because one variety of professional services is not a perfect substitute for another, in general. In the special case in which it is – that is $\rho = 1$ – the price of a variety is then directly proportional to its marginal value.

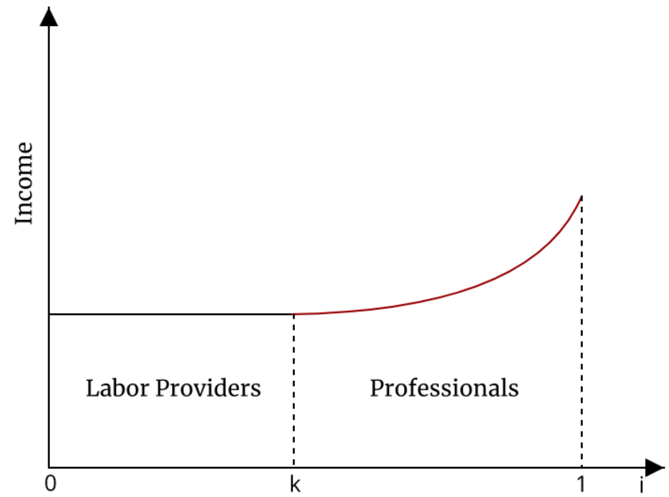

Figure 1: The Equilibrium Income Distribution

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

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

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 provide labor 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.[17]

We find that k is determined by the following market clearing condition for the alternative good, where $k_0 = \frac{Bc}{A+Bc}$.

[17]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.

Proposition 1. *The identity of the marginal professional, k, is determined by*

$$\mu^{\frac{1}{1-\widehat{\rho}}} H_k^{\frac{\rho-\widehat{\rho}}{1-\widehat{\rho}}} h_k^{\frac{-\rho}{1-\widehat{\rho}}} \times (A/B+c)^{\frac{\widehat{\rho}}{1-\widehat{\rho}}} = k - k_0. \tag{17}$$

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. ∎

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 (17) 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.[18]

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.[19] 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

[18]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.

[19]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. 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 $k=1$.

economy can support only the agent with the greatest 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 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 production 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. Considering that agents with higher levels of human capital charge higher prices and earn greater income, this observation suggest 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 be increased. This intuition is confirmed by the following proposition.

Proposition 2. *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 she should expect the coefficient be positive. Observe, however, that while the changes to the super high earning follows the pattern of Proposition 2, their relative positions might not be affected much by a rise in B. That 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. However, this result, as concerned with comparison within super earners, is not in conflict with the observation that the top 1 or top 0.1 percent earners are reaping ever a higher proportion of incomes relative to the rest.

In the above intuitive discussion, we expect the a rise in B should cause the prices of all professional services to fall because it intensifies competition between them. This fall is indeed confirmed by the following proposition.

Proposition 3. *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 go down, then that means 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 2 – all professionals reap a net gain from the rise in B. If, on the contrary, 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 thus shows that the movement in the cutoff k depends on 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 marginal professional is better off and someone with lesser human capital becomes the new marginal professional – that is, the cutoff k falls. In contrast, if the negative effect due to stronger competition dominates, then the present marginal professional would be worse off if she sticks to the professional occupation. As a result, she exits the profession and someone with more human capital becomes the new marginal professional – that is, the cutoff k rises. The following proposition determines the sufficient and necessary condition for either scenario to arise in equilibrium.

Proposition 4. *There exists a finite $\overline{\mu} > 0$ such that $\frac{dk}{dB} > 0$ if and only if*

$$\widehat{\rho} \leq k_0 \text{ or } \mu < \overline{\mu}. \tag{18}$$

Proof. We relegate the proof to B.4. ∎

According to Proposition 4, under condition (18), with a rise in B the negative effect due to stronger competition dominates the positive effect due to capacity enlargement for the lower-end professionals, and, as a result, they are squeezed out of the professional occupation and enter the alternative occupation. Condition (18) requires either that the substitutability between professional services and the alternative good $\widehat{\rho}$ is not too big or that the relative importance of the latter to the former, μ, is not too high. To intuitively understand why under this condition a rise in B lowers the the prices of services by such a big scale, observe that a fall in the service prices shifts agents' demand from the alternative good to the professional services. This provides a channel to absorb the increase in the supply of the services caused by the rise in B. The service prices will fall by a large amount, as a result, if this 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 much substitutable to the alternative good, then consuming more of the former diminish little of agents' valuation of the latter; indeed, they value the alternative good even more if $\widehat{\rho} < 0$, namely, if professional services are complementary to it. 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 to the alternative good, one unit less of the alternative good can absorb – that is, substitute – only μ units of professional services.

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 1. *Under condition (18) 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 4, 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 (18), this countervailing force dominates, lowering some workers' incomes[20] and inducing exit from the affected occupations.

While lower-end professionals clearly lose under condition (18), 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

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

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

[20] One caveat is that these workers' welfare might still be higher, despite their income measured with the constant real good (i.e. the alternative good) is reduced, because all the professional services are cheaper (Proposition 3).

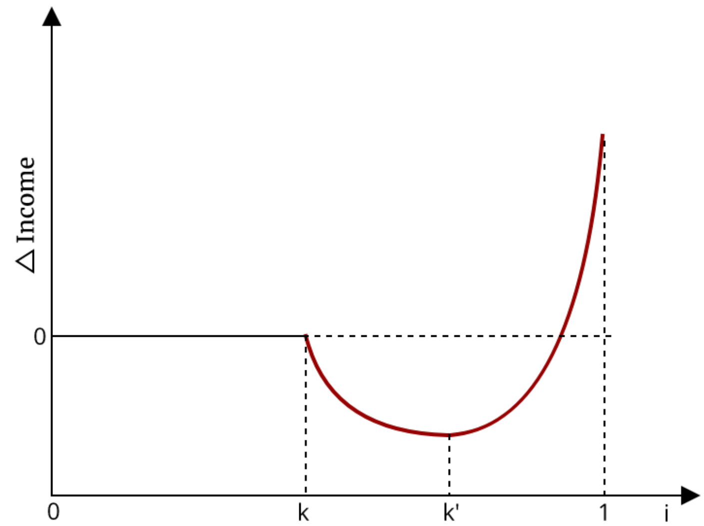

Figure 2: Income growth due to an expansion in B.

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

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

Proposition 5. *Assume that $\Omega(\rho)\cdot A/(A+Bc)<1$ and $\widehat{\rho}\geq 0$. If $h_1>1$ and $h_1\geq \xi\cdot h(f(k_0, D\cdot \xi^{\frac{\rho}{1-\widehat{\rho}}}))$, then $dm_1/dB>0$.*

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

Note that this proposition is concerned with only two points along the function $h(i)$, namely at $i=1$ and $i=f(k_0, D\cdot\xi^{\frac{\rho}{1-\widehat{\rho}}})$, and hence it can be satisfied by any distribution of human capital in which $h(1)$ is sufficiently large. When this proposition holds, the top professionals gain in net from an increase in B, while under condition (18), professionals at the bottom of the distribution suffers a net loss. Hence, under the conditions of Proposition 5 and condition (18), 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,

$$\frac{\rho}{1-\rho}\frac{A}{c} > B. \tag{19}$$

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. 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 1 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 2 still holds, that is, an increases in B raises inequality within the professional occupation.

Proposition 6. *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 3 also holds. A rise in B intensifies competition between different varieties of professional services, causing their prices to fall.

Proposition 7. *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 4 to hold under a condition that is less strict relative to (18). 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 its negative effect, namely that it leads to stronger competition between professionals. 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 1), 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 5 holds, although the exact conditions that describe what the term "high enough" means will be different.

5 A Comparison to Technological Change that Reduces Marginal Cost

As we have argued throughout, and as highlighted by Corollary 1 in Subsection 4.2, technology that increases the scale of operation is special in that although it technologically complements professionals, 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 (18). To appreciate the specialty of this type of technological change, we compare it to another type of technological change that reduces the marginal cost c of professionals' production. Of course, in reality technological changes may alter production and consumption behavior in more than one way. For example, the internet not only increases the size of the market for some workers, it also reduces communication costs and thereby reduces the marginal cost of production for occupations that are communication intensive.

Consider a technological change that reduces the 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 8. *$\frac{dk}{d(-c)} < 0$ 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 opposite to that of an increase in the scale of operation B under condition (18). 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 (18) 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 under an increase in B. We note that the effect of a marginal cost reduction is intuitive – that is, a technologically favorable change typically benefits the affected workers and draws people into the affected occupation, which is of course not true for a rise in B due to the competition effect.

6 Empirical Evidence

In this section we take the advent and growth of new ICT (primarily the internet) over the past 25 years as reflecting a rise in the scale of operation B, and we highlight direct links between our formal model predictions and some basic facts regarding the link between ICT growth and rising within-occupation inequality.

6.1 Measuring an Occupation's "Exposure" to the Internet

As a starting point, we seek to link the model parameter B, reflecting the scale of operation associated with some occupation, to an empirical counterpart. To do this, we note that the extent to which an occupation experienced a rise in its scale of operation during the internet era will depend on how occupational output was impacted by the internet-induced growth in market access. To formalize this idea, we construct a measure of the extent to which the output of each of 330 U.S. occupations was associated with internet sales in 2000 and 2010, setting 1990 to zero for all occupations (since this was prior to the advent of the modern internet). We denote the measure as B^{int} in order to link it conceptually to the scale of operation that was the focus of the theoretical section. We define the measure in the following way:

$$B_{it}^{int} = \sum_j \left(IntShr_{jt} \times OccShr_{ij,1990} \right) \tag{20}$$

where $IntShr_{jt}$ is the share of industry j sales in year t that was made over the internet and $OccShr_{ij,1990}$ is the share of occupation i's total hours employed in industry j in 1990.[21]

Thus, the latter term in (20) reflects the importance of each industry, in terms of labor hours, to each occupation in a period in which the internet was absent. Note that we use this *pre-period* occupational structure in order to avoid incorporating effects due to endogenous changes in the composition of occupations that were themselves caused by the internet. The former term then

[21] The industry internet sales data come from Census' E-Stats database, which provides the data at the two- and three-digit North American Industry Classification System (NAICS) level. We then concord these to the Ind1990 classification used in the CPS using a straightforward concordance provided by Census. One nuance is that some of the sales data is classified under the industry "E-Merchants" (NAICS 4541) by product, in categories such as Books and Magazines, Music and Videos, etc. We therefore match these to the relevant Ind1990 industries manually.

	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 according to (20).

captures the extent to which firms within each industry sell their output over the internet.[22] Table 1 lists the top 10 (left column) and bottom 10 (right column) occupations in terms of their exposure to internet sales according to this measure.

6.2 Decomposing Wage Inequality

Next, we seek to document changes in wage inequality within occupations over this time period. To do this, we decompose aggregate log wage dispersion into within- and between-occupation components separately for 1990, 2000 and 2010. Following this, we will use our internet exposure measure to further characterize this decomposition. Formally, we calculate the following:

$$\frac{1}{N_t}\sum_i (w_{it}-\bar{w}_t)^2 = \frac{1}{N_t}\sum_l \sum_{i\in l}(w_{it}-\bar{w}_{lt})^2 + \frac{1}{N_t}\sum_l N_{lt}(\bar{w}_{lt}-\bar{w}_t)^2 \tag{21}$$

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 2 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. We can decompose total log wage dispersion further by noting that the within term in (21) is the sum across individual occupations, and so the

[22] 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.

	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 2: Wage dispersion within and between occupations, 1990-2010, as defined in (21). "Top 10% Most Exposed" are the top 10 percent most exposed occupations according to (20) and "Other" are all other occupations. The table indicates that 40 percent of the rise in aggregate wage inequality was due to a rise in within-occupation inequality. Further, 65 percent of this within variation was due to the top 10 percent most-exposed occupations to the internet.

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, as defined by our measure (20). To illustrate this more clearly, the last four columns of Table 2 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 (20) 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 top 10 percent of occupations that were most exposed to internet sales.

Finally, we note that we have also performed this decomposition in wage inequality using log *residual* income. In this case, the income variable is first "cleaned" of variation due to age, age squared, sex and education level by running a regression in which the log wage is on the left hand side and these variables are all on the right hand side. The residual from this regression is then used as the wage measure going forward, and as such is described as being "cleaned" of variation that may independently be contributing to our outcomes of interest. See Appendix A, which indicates that the results are quite similar, suggesting that the effects of these variables are not problematic for our analysis.

6.3 Linking the Facts to the Model

Figures 3 and 4 bring the formal decomposition of wage inequality presented in Table 2 into graphical form, highlighting the relative contribution of different groups to aggregate wage inequality. In this section we show that the facts reflected in these Figures are consistent with the primary results from the model. We then introduce additional facts that we link to model predictions. We highlight three specific links between model and facts.

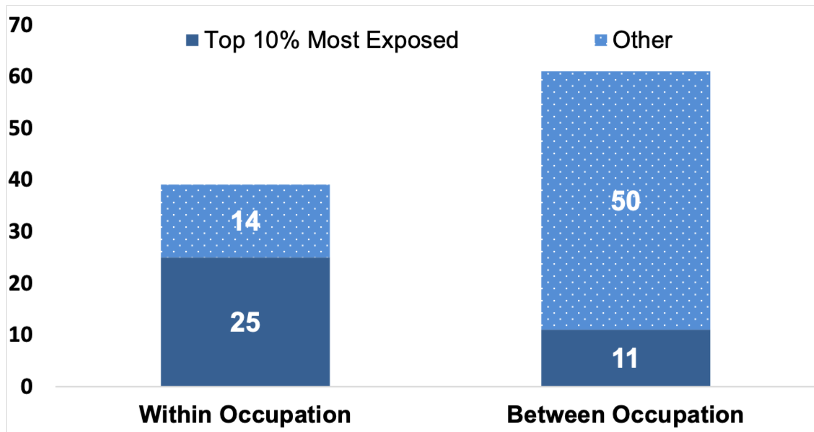

Figure 3: Contributions to Growth in U.S. Wage Inequality, 1990-2010. "Top 10% Most Exposed" is defined as the top 10 percent of occupations most exposed to the internet according to (20) and "Other" are all other occupations. 40 percent of the growth in U.S. wage inequality occurred within occupations, most of which was due to occupations most exposed to the internet.

1. **Propositions 2 and 6 predict that ICT is a driving force for increased inequality within occupations.** Consistent with this, Figure 3 indicates that 40 percent of the rise in aggregate wage inequality over the period 1990-2010 occurred *within* occupations. In addition, nearly two-thirds of the increase is due to the top ten percent of occupations that were most exposed to rising sales over the internet, based on measure (20).

2. **Corollary 1 predicts that ICT innovation leads to income declines for low-end workers within affected occupations.** Figure 4 documents the change in wages over this period at different parts of the initial (1990) wage distribution, again separating out occupations that were most exposed to the internet. Consistent with Corollary 1, Figure 2 indicates that the bottom 40 percent of workers (by initial 1990 wage) in the most exposed occupations saw wage declines over the 1990-2010 period, while lower-end workers in unexposed occupations did not.

3. **Propositions 3 and 7 predict that ICT innovation lowers the price of goods sold using the new technology.** Consistent with this, a growing body of literature has consistently found that prices on the internet are lower than otherwise identical items sold elsewhere, likely due to increased competition arising from reduced search costs (see, e.g., Brynjolfsson and Smith (2000); Zettelmeyer et al. (2006); or Cavallo (2018)). For instance, Brown and Goolsbee (2002) find that the internet reduced term life insurance prices by 8 to 15 percent.

To summarize, 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 innovations in ICT. In fact, the first fact indicates that over a quarter of the recent rise in *aggregate* wage inequality can be explained by these types of occupations. The second fact indicates that the least

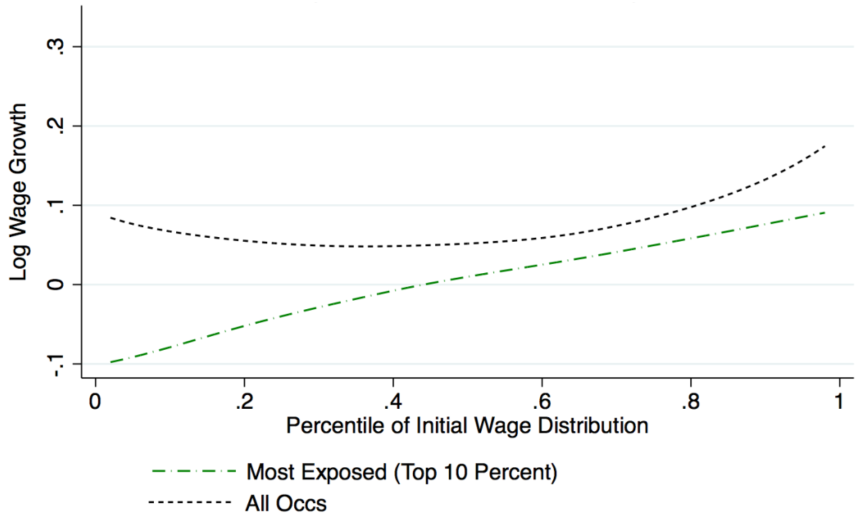

Figure 4: Occupation exposure to the internet, according to (20), and log wage growth, 1990-2010. The most exposed occupations saw negative wage growth among initially low-wage workers, and rising wage growth at higher initial wages. Other occupations saw positive wage growth throughout the distribution of initial wages.

skilled workers in the most exposed occupations have, on average, seen wage declines over the last two decades, consistent with a rise in competition among workers for these jobs. Finally, the third fact also highlights a potential impact of increased competition on prices in markets in which the internet is a key driver of sales. Our model generates predictions that are consistent with these key economic outcomes.

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 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

Acemoglu D, Autor D. 2011a. Skills, Tasks and Technologies: Implications for Employment and Earnings. In Ashenfelter O, Card D (eds.) *Handbook of Labor Economics*, volume 4, chapter 12. Elsevier, 1043–1171.

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

Anderson SP, Renault R. 1999. Pricing, product diversity, and search costs: A bertrand-chamberlin-diamond model. *The RAND Journal of Economics* : 719–735.

Armstrong M, Vickers J, Zhou J. 2009. Prominence and consumer search. *The RAND Journal of Economics* **40**: 209–233.

Autor D, Dorn D, Katz LF, Patterson C, Van Reenen J, et al. 2017. *The fall of the labor share and the rise of superstar firms.* National Bureau of Economic Research.

Autor DH, Levy F, Murnane RJ. 2003. The Skill Content Of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics* **118**: 1279–1333.

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

Bar-Isaac H, Caruana G, Cuñat V. 2012. Search, design, and market structure. *American Economic Review* **102**: 1140–60.

Beaudry P, Doms M, Lewis E. 2010. Should the Personal Computer Be Considered a Technological Revolution? Evidence from U.S. Metropolitan Areas. *Journal of Political Economy* **118**: 988 – 1036.

Bessen JE. 2017. Information technology and industry concentration .

Brown JR, Goolsbee A. 2002. Does the internet make markets more competitive? evidence from the life insurance industry. *Journal of political economy* **110**: 481–507.

Brynjolfsson E, Smith MD. 2000. Frictionless commerce? a comparison of internet and conventional retailers. *Management science* **46**: 563–585.

Cachon GP, Terwiesch C, Xu Y. 2008. On the effects of consumer search and firm entry in a multiproduct competitive market. *Marketing Science* **27**: 461–473.

Cavallo A. 2018. More amazon effects: Online competition and pricing behaviors. Technical report, National Bureau of Economic Research.

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

Chen Y, He C. 2011. Paid placement: Advertising and search on the internet. *The Economic Journal* **121**: F309–F328.

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

Egger H, Kreickemeier U. 2012. Fairness, Trade, and Inequality. *Journal of International Economics* **86**: 184–196.

Firpo S, Fortin NM, Lemieux T. 2012. Occupational Tasks and Changes in the Wage Structure. Textos para discussão 284, Escola de Economia de São Paulo, Getulio Vargas Foundation (Brazil).

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

Garicano L, Rossi-Hansberg E. 2004. Inequality and the Organization of Knowledge. *American Economic Review* : 197–202.

Garicano L, Rossi-Hansberg E. 2006. The knowledge economy at the turn of the twentieth century: the emergence of hierarchies. *Journal of the European Economic Association* **4**: 396–403.

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

Hornstein A, Krusell P, Violante GL. 2011. Frictional wage dispersion in search models: A quantitative assessment. *American Economic Review* **101**: 2873–98.

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

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

Kim C, Sakamoto A. 2008. The rise of intra-occupational wage inequality in the united states, 1983 to 2002. *American Sociological Review* **73**: 129–157.

Koulayev S. 2014. Search for differentiated products: identification and estimation. *The RAND Journal of Economics* **45**: 553–575.

Krugman PR. 1979. Increasing returns, monopolistic competition, and international trade. *Journal of international Economics* **9**: 469–479.

Kugler M, Verhoogen E. 2012. Prices, Plant Size, and Product Quality. *The Review of Economic Studies* **79**: 307–339.

Larson N. 2013. Niche products, generic products, and consumer search. *Economic Theory* **52**: 793–832.

Lucas Jr RE. 1978. On the Size Distribution of Business Firms. *The Bell Journal of Economics* : 508–523.

Melitz MJ. 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica* **71**: 1695–1725.

Melitz MJ, Ottaviano GI. 2008. Market size, trade, and productivity. *The review of economic studies* **75**: 295–316.

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

Mouw T, Kalleberg AL. 2010. Occupations and the structure of wage inequality in the united states, 1980s to 2000s. *American Sociological Review* **75**: 402–431.

Neal D, Rosen S. 2000. Theories of the Distribution of Earnings. In Atkinson A, Bourguignon F (eds.) *Handbook of Income Distribution*, volume 1 of *Handbook of Income Distribution*, chapter 7. Elsevier, 379–427.

Rosen S. 1981. The Economics of Superstars. *American Economic Review* **71**: 845–58.

Rosen S. 1983. Specialization and Human Capital. *Journal of Labor Economics* **1**: 43–49.

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

Saint-Paul G. 2007. Knowledge Hierarchies in the Labor Market. *Journal of Economic Theory* **137**: 104–126.

Stahl DO. 1989. Oligopolistic pricing with sequential consumer search. *The American Economic Review* : 700–712.

Tinbergen J. 1974. Substitution of graduate by other labour*. *Kyklos* **27**: 217–226.

Zettelmeyer F, Morton FS, Silva-Risso J. 2006. How the internet lowers prices: Evidence from matched survey and automobile transaction data. *Journal of marketing research* **43**: 168–181.

Zhou J. 2014. Multiproduct search and the joint search effect. *American Economic Review* **104**: 2918–39.

APPENDIX

A Wage Inequality Using Variation in the Log Residual Wage

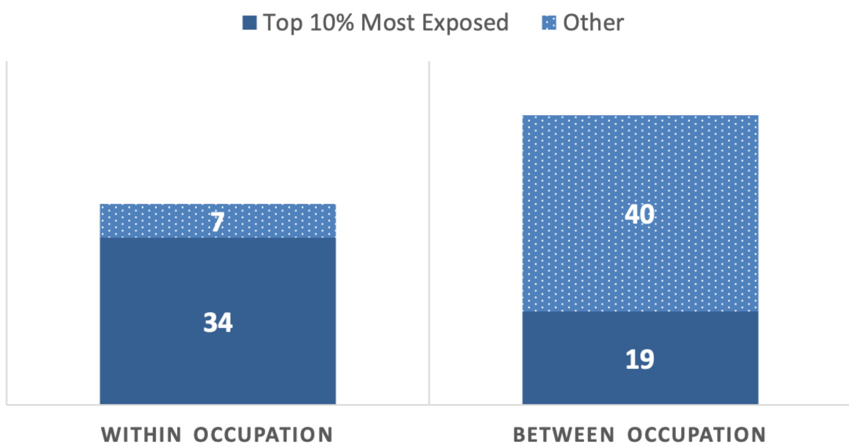

Figure 5: Contributions to Growth in U.S. Wage Inequality, 1990-2010. “Top 10% Most Exposed” is defined as the top 10 percent of occupations most exposed to the internet according to our measure defined in (20), and “Other” are all others. Here, the wage is defined as the residual of a regression of log wage on age, age squared, sex and education level.

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

$$P = \left(\frac{Mf(P,\mu)}{B}\right)^{1-\rho} H_k^{\rho-1}. \tag{22}$$

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

$$M = BPH_k \times \left[1+\mu^{\frac{1}{1-\widehat{\rho}}}P^{\frac{\widehat{\rho}}{1-\widehat{\rho}}}\right]. \tag{23}$$

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:

$$BH_k(\mu P)^{\frac{1}{1-\widehat{\rho}}} = (A+Bc)(k-k_0). \tag{24}$$

This equation contains P, an endogenous variable. To determine k, we exploit an additional connection between P and k, as follows. Using (22) 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,

$$PH_k^{1-\rho}h_k^{\rho} = \frac{A}{B}+c. \tag{25}$$

Solving for P, substituting it into (24) and rearranging, we arrive at equation (17) 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,

$$\frac{d\log m_i}{dB} = \frac{d\log\widetilde{m}_i}{dB}\times\left(1+\frac{Bc}{m_i}\right)-\frac{c}{m_i}. \tag{26}$$

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

$$\frac{d\log\widetilde{m}_i}{dB} = \frac{c}{Bc+A} - \rho\left[\log h_k\right]'_k\frac{dk}{dB} \tag{27}$$

and is independent of m_i. Hence, from (26), $\left[\frac{d\log m_i}{dB}\right]'_{m_i} = \frac{d\log \widetilde{m}_i}{dB} \times \left[1+\frac{Bc}{m_i}\right]'_{m_i} - \left[\frac{c}{m_i}\right]'_{m_i} = \frac{d\log \widetilde{m}_i}{dB} \times \frac{-Bc}{m_i^2} + \frac{c}{m_i^2}$. It follows that $\left[\frac{d\log m_i}{dB}\right]'_{m_i} > 0$ if and only if:

$$\frac{d\log \widetilde{m}_i}{dB} < \frac{1}{B}. \tag{28}$$

Using equation (27), this equality is equivalent to $\left[\frac{c}{Bc+A} - \rho\left[\log h_k\right]'_k \frac{dk}{dB}\right] B < 1 \Leftrightarrow$

$$B\rho\left[\log h_k\right]'_k \left(-\frac{dk}{dB}\right) < \frac{A}{Bc+A}. \tag{29}$$

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 (29), and the inequality becomes:

$$\rho\left[\log h_k\right]'_k k_0 \left(-\frac{dk}{dk_0}\right) < 1. \tag{30}$$

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 (17) to find an upper bound for $-\frac{dk}{dk_0}$. Because $A/B + c = c/k_0$, this condition can be rearranged as

$$\mu^{\frac{1}{1-\widehat{\rho}}} H_k^{\frac{\rho-\widehat{\rho}}{1-\widehat{\rho}}} h_k^{\frac{-\rho}{1-\widehat{\rho}}} = (k-k_0)\,(\frac{k_0}{c})^{\frac{\widehat{\rho}}{1-\widehat{\rho}}}. \tag{31}$$

Taking the logarithm on both sides we have:

$$-\frac{dk}{dk_0} = \frac{\frac{\widehat{\rho}k-k_0}{(1-\widehat{\rho})k_0(k-k_0)}}{\frac{1}{k-k_0} + \frac{\rho}{1-\widehat{\rho}}\left[\log h_k\right]'_k - \frac{\rho-\widehat{\rho}}{1-\widehat{\rho}}\left[\log H_k\right]'_k}. \tag{32}$$

By assumption, $\left[\log h_k\right]'_k > 0$. Since (14) implies that $H_x^{\rho} = \int_x^1 h_i^{\rho}$ decreases with x, we have $\left[\log H_k\right]'_k < 0$. Furthermore, by assumption $\frac{\rho-\widehat{\rho}}{1-\widehat{\rho}} > 0$. It follows that all three terms in the denominator of (32) are positive. $-\frac{dk}{dk_0} > 0$ if and only if the numerator is positive. If it is positive then we have

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

As a result, if $-\frac{dk}{dk_0} > 0$, inequality (30) follows from $\rho\left[\log h_k\right]'_k k_0 \times \frac{\widehat{\rho}k-k_0}{k_0(k-k_0)\times\rho\left[\log h_k\right]'_k} \leq 1$, which is equivalent to $\widehat{\rho}k - k_0 \leq k - k_0$, 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

$$\frac{d\log m_i}{dB} < \frac{1}{B}. \tag{33}$$

In the proof of Proposition 2, we have proved that $\frac{d\log \widetilde{m}_i}{dB} < \frac{1}{B}$ in (28). 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 (26), 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 (32) (see the proof of Proposition 2), holds true if and only if

$$\widehat{\rho}k < k_0. \tag{34}$$

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 (31) (see the proof of Proposition 2), 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 (34) holds – if and only if $g(k) < g(k_0/\widehat{\rho})$, that is, $\mu < g(k_0/\widehat{\rho})$. Define $\overline{\mu} := g(k_0/\widehat{\rho})$ and we complete the proof. Furthermore, if $\widehat{\rho} > k_0$ and hence $k_0/\widehat{\rho} < 1$, we have $g(k_0/\widehat{\rho})$ is finite. ∎

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 2. ∎

B.6 Proof of Proposition 5

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

Lemma 1. *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*

$$\frac{h_i}{h_k} > \left(\frac{1}{1 - \Omega(\rho) \cdot A/(A + Bc)}\right)^{\frac{1}{\rho}}. \tag{35}$$

Proof of Lemma 1

Proof. By (16),

$$\frac{dm_i}{dB} = h_i^\rho/h_k^\rho \cdot [c - (A+Bc)\cdot\rho\cdot(\log h_k)' \cdot \frac{dk}{dB}] - c. \tag{36}$$

The identity of the marginal professional, k, is determined by equation (44). 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,

$$\frac{dk}{dB} < \frac{Ac/(A+Bc)^2}{1+\rho(\log h_k)'(k-k_0)}. \tag{37}$$

By (36), $\frac{dm_i}{dB} > 0$ if

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

With an upper bound of $\frac{dk}{dB}$ given by (37), this inequality follows from: $h_i^\rho/h_k^\rho \cdot [c - (A+Bc)\cdot\rho\cdot(\log h_k)'\cdot\frac{Ac/(A+Bc)^2}{1+\rho(\log h_k)'(k-k_0)}] > c \Leftrightarrow$

$$h_i^\rho/h_k^\rho \cdot [1 - \frac{A}{A+Bc}\cdot\frac{\rho\cdot(\log h_k)'}{1+\rho(\log h_k)'(k-k_0)}] > 1, \tag{39}$$

which is equivalent to (35). ∎

Condition (35), 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 2. *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

$$k - k_0 < D(\frac{h_1}{h_k})^{\frac{\rho}{1-\widehat{\rho}}}(1-k)^{\frac{\rho-\widehat{\rho}}{\rho(1-\widehat{\rho})}}. \tag{40}$$

Proof: k is determined by equation (44), 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 (40).

Step 2:

$$k < f(k_0, D \cdot (\frac{h_1}{h_k})^{\frac{\rho}{1-\widehat{\rho}}}). \tag{41}$$

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 (40) 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 (41) 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 2, we let $\widetilde{m}_i := m_i + Bc = p_{i^*} B \times h_i^\rho / h_{i^*}^\rho$ be the revenue. Then

$$\frac{d \log \widetilde{m}_i}{dB} = \frac{1}{B} - \rho \frac{h_{i^*}'}{h_{i^*}} \frac{di^*}{dB}. \tag{42}$$

The proof of Proposition 2 has shown that $\left[\frac{d \log m_i}{dB}\right]_{m_i}' > 0$ if and only if:

$$\frac{d \log \widetilde{m}_i}{dB} < \frac{1}{B}, \tag{43}$$

which, with (42), 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 3, $\frac{dp_i}{dB} < 0$ if $\frac{d\log \widetilde{m}_i}{dB} < \frac{1}{B}$, which is proved as inequality (43) in the proof of Proposition 6. ∎

B.9 Proof of Proposition 8

Proof. Equilibrium condition (17) is equivalent to

$$\mu^{\frac{1}{1-\widehat{\rho}}} H_k^{\frac{\rho-\widehat{\rho}}{1-\widehat{\rho}}} h_k^{\frac{-\rho}{1-\widehat{\rho}}} = (A/B + c)^{\frac{-\widehat{\rho}}{1-\widehat{\rho}}} \left[k - Bc/\left(A+Bc\right)\right]. \tag{44}$$

Use the implict function theorem and recall that $k_0 = Bc/\left(A+Bc\right)$, 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 4, 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 4 and show that $\left[\frac{d\log m_i}{dc}\right]_{m_i}' > 0$. Let $\widetilde{m}_i := m_i + Bc$. Then

$$\frac{d\log m_i}{dc} = \frac{d\log \widetilde{m}_i}{dc} \times \frac{\widetilde{m}_i}{m_i} - \frac{B}{m_i}. \tag{45}$$

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

$$\frac{d\log \widetilde{m}_i}{dc} = \frac{B}{Bc + A} - \rho\frac{h_k'}{h_k}\frac{dk}{dc}. \tag{46}$$

Observe that $\frac{d\log \widetilde{m}_i}{dc}$ is independent of m_i. Hence, from (45), $\left[\frac{d\log m_i}{dB}\right]_{m_i}' = \frac{d\log \widetilde{m}_i}{dc} \times \left[\frac{m_i+Bc}{m_i}\right]_{m_i}' - \left[\frac{B}{m_i}\right]_{m_i}' = \frac{d\log \widetilde{m}_i}{dc} \times \frac{-Bc}{m_i^2} + \frac{B}{m_i^2}$. 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 (27), is equivalent to $-\frac{Bc}{Bc+A} + \rho c\frac{h_k'}{h_k}\frac{dk}{dc} + 1 > 0$, which certainly holds true as $\frac{dk}{dc} > 0$. ∎