

The Spillover Effects of Top Income Inequality*

Jeffrey Clemens[†] Joshua D. Gottlieb[‡] David Hémous[§]
Morten Olsen[¶]

November 2017
Preliminary and incomplete

Abstract

Top income inequality in the United States has increased considerably within occupations as diverse as bankers, managers, doctors, lawyers and scientists. The breadth of this phenomenon has led to a search for a common explanation. We show instead that increases in income inequality originating within a few occupations can “spill over” into others, driving broader changes in income inequality. We develop an assignment model where *consumers* with heterogeneous income buy services from *doctors* with heterogeneous ability. In equilibrium the highest-earning consumers match with the highest-ability doctors. Increases in income inequality among the consumers feed directly into the doctors’ income inequality. To test our theory, we identify occupations for which our consumption-driven theory predicts spillovers and occupations for which it does not. Using a Bartik-style instrument, we show that an increase in general income inequality causes higher income inequality for doctors, dentists and real estate agents,

*Morten Olsen gratefully acknowledges the financial support of the European Commission under the Marie Curie Research Fellowship program (Grant Agreement PCIG11-GA-2012-321693) and the Spanish Ministry of Economy and Competitiveness (Project ref: ECO2012-38134). Gottlieb acknowledges support from the Social Sciences and Humanities Research Council of Canada and the Upjohn Institute for Employment Research. We are grateful to numerous people and organizations for their generosity in sharing data and facilitating data access: Daniel Polsky, Michael Levin, Leemore Dafny, Nicole Fortin, Cecilia Ganduglia-Cazaban, the UTSPH/BCBSTX Payment Systems and Policies Research Program at the University of Texas School of Public Health, CIVHC, Statistics Canada, SSHRC, and the staff at the British Columbia Interuniversity Research Data Centre at UBC. We are deeply indebted to Jeff Hicks for truly exceptional research assistance, and thank Kenta Baron-Furuyama, Innessa Colaiacovo, Erin May, and Andrew Vogt for outstanding research assistance as well. Finally, we thank Paul Beaudry, David Green, Jason Sutherland, and numerous seminar participants for valuable comments.

[†]University of California, San Diego and NBER

[‡]University of British Columbia and NBER

[§]University of Zürich and CEPR

[¶]University of Copenhagen

and in fact accounts for most of the increases in inequality within these occupations. Physician pricing and insurance network data support our mechanism.

JEL: D31; J24; J31; O15

Keywords: Income inequality, Assignment model, Occupational inequality, Superstars

1 Introduction

Since the 1980s the share of total earnings going to the top of the income distribution has increased considerably. Even within the top income inequality has increased: A higher share of top earnings accrue to the very high earners. Figure 1 shows that this pattern holds within specific occupations so that the overall growth of top income inequality is not simply due to the divergence between bankers and lawyers, or between programmers and physicians (Bakija, Cole, and Heim, 2012). At first glance, this broad pattern suggests that any plausible explanation for rising inequality—whether it be globalization, deregulation, changes to the tax structure, or technological change—would have to apply to occupations as diverse as financial managers, doctors, and CEOs (Kaplan and Rauh, 2013). We argue that this need not be the case because inequality across occupations is linked. We show that exogenous increases in income inequality within one occupation “spill over” into others through the former’s consumption. This drives up income inequality for a broader set of occupations than those affected by the initial shock.

We present a model where changes in within-occupation income inequality propagate to other occupations through consumption. We do not rely on competition for skill in the broader labor market. We develop an assignment model where *consumers* with heterogeneous income buy the services of *doctors* with heterogeneous ability. In equilibrium the highest-earning consumers match with the highest-ability doctors. The prices necessary to rationalize this match cause income inequality among the consumers to feed directly into income inequality among doctors.

Two conditions on the services provided by doctors are necessary for the equilibrium to feature an assignment mechanism and hence income inequality spillovers: *heterogeneity* and *non-divisibility* in output. Non-divisibility means that one high-ability doctor is not the same as two decent-ability doctors. We focus on physicians, dentists and real estate agents, occupations that meet these conditions, and contrast them with occupations that do not. Using Census data from 1980–2014, we find that an increase in general income inequality causes an increase in inequality for these occupations, with a spillover elasticity ranging from 0.5 to 2.7. These occupations are important within the top one percent of the income distribution—in fact Physicians are the most common Census occupation in the top 1%.

In support of our model’s assortative matching mechanism, we present empirical evidence on how health care spending and physician prices relate to household income. Using both a nationally representative survey, and more detailed medical claims data for a convenience sample of patients, we find that use of more expensive providers drives a large share of the

health spending Engel curve. Consumers earning ten percent more see physicians who are 1.9 percent more expensive, while spending 2.3 percent more on health care overall.

Our baseline model considers occupations of heterogeneous ability where production is not scalable; that is, no mechanism exists that would allow the more talented to scale up output. For these occupations, when consumption is non-divisible, the income distribution is tightly linked to that of the general population. Specifically, *consumers* of heterogeneous ability produce a homogeneous product in quantity proportional to their skill level. Besides this homogeneous product, each consumer purchases the services of one *doctor*. Doctors also have heterogeneous ability but their ability translates proportionately into the quality of the services they provide and not the quantity. All doctors serve the same number of patients. Consumers and doctors both have Pareto-distributed ability but with possibly different parameters.

Equilibrium in this model is an assignment function with positive assortative matching: the highest-ability consumers match with the highest-ability doctors. An exogenous mean-preserving spread in the income inequality of consumers increases the number of high-earning consumers. This increases the demand for the best doctors and increases top income inequality among doctors as well. In fact, in the special case of Cobb-Douglas utility, top income inequality of doctors is entirely driven by the earnings distribution of consumers and is independent of the underlying ability distribution of doctors.

We extend the model in three directions. First, we allow for occupational mobility at the top: high-ability doctors can choose to be high-ability consumers and vice versa. Since changes in top income inequality for consumers completely translate into changes in top income inequality for doctors when there is no occupational mobility, allowing occupational mobility has no impact on doctors' inequality; the two settings are observationally equivalent. Second, we consider two regions in one nation that differ only in the top income inequality of consumers. We allow patients to import their medical services. We show that top income inequality among doctors for each region must follow consumer top income in the most unequal region. This distinction will be important for the empirical test: When a service is non-tradable, spillover effects will happen at the local level, whereas for tradable services the spillover effect will happen at the national level. Finally, we let doctors move across regions and show that the most unequal region will attract the most able doctors. But, as in the baseline model, doctors' inequality is determined by general inequality in the region where they eventually live. Hence the observed top income inequality of doctors is the same whether or not they can move.

We use four broad strategies to test different components of this model. First, we establish the unsurprising fact that higher-income patients visit higher-priced physicians—a necessary condition required for spillovers to arise. Second, we show the key causal result: an increase in local income inequality among non-physicians leads to increased inequality among physicians. To establish this causal link, we use a Bartik (1991)-style instrument. We construct a weighted average of nationwide inequality for the 20 occupations that are the most represented in the top ten percent nationwide (excluding the occupation of interest). The weights correspond to the relative importance of each occupation in each labor market area at the beginning of our sample. In other words, we only exploit the changes in labor market income inequality that arise from the occupational distribution in 1980 combined with the nationwide trends in occupation-specific inequality. This weighted average serves as our instrument for general inequality in the area in question.

Using this instrument, we find a very clear distinction between the first group and other occupations. An increase in general income inequality at the local level causes an increase in inequality for physicians, dentists and real estate agents, who operate in local markets. The parameter estimates suggest that the majority of the increase in income inequality for these occupations can be explained by increases in general income inequality.

Our third empirical finding uses the theory to predict which occupations will feature spillovers: those with non-divisibility in output. Furthermore, since we focus on geographical variation across the United States, our estimation methodology will only pick up spillover effects if they are local—that is, if workers mostly service local clients. We classify occupations into two groups: Those that meet these conditions (such as physicians, dentists and real estate agents) and those that do not (such as financial managers, college professors and secretaries.) Using the same empirical strategy, we find that an increase in general income inequality (excluding the occupation of interest) is positively correlated with an increase in inequality for occupations in the first group, such as physicians. On the other hand, as our theory predicts, we find that *local* general income inequality does *not* spill over to financial managers, college professors and secretaries.

Fourth, our model proposes a specific mechanism for transmitting inequality from the general population to private physicians: price inequality and assortative matching. Physicians in more unequal areas should charge unequal prices. We use detailed physician claims data from three states to directly examine this mechanism. Since actual physician payments in the United States reflect the structure of insurance plan networks, we also examine inequality in these network sizes. Both types of data support the mechanism we propose:

pricing inequality is increasing in areas with growing inequality, and network size is more heterogeneous in more unequal markets.

The increase in top income inequality has inspired substantial scholarship (among many others, see Piketty and Saez, 2003, and Atkinson, Piketty and Saez, 2011). This literature has established that at the top, the income distribution is well-described by a Pareto distribution (see Guvenen, Karahan, Ozkan, and Song, 2015, for some of the most recent evidence, and Pareto, 1896, for the earliest). Further, Jones and Kim (2014) show that the increase in top income inequality is linked with a fattening of the right tail of the income distribution, which corresponds to a decrease in the shape parameter of the Pareto distribution. This literature is related to, but distinct from, the large literature on skill-biased technological change and income inequality which seeks to explain changes in income inequality throughout the income distribution and primarily *across* occupations (Goldin and Katz, 2010; Acemoglu and Autor, 2011.)

More specifically, our paper builds on the “superstars” literature originating with Rosen (1981), who explains how small differences in talent may lead to large differences in income. The key element in his model is an indivisibility of consumption result which arises from a fixed cost in consumption per unit of quantity. This leads to a “many-to-one” assignment problem as each consumer only consumes from one performer (singer, comedian, etc.), but each performer can serve a large market (see also Sattinger, 1993).¹ In that framework, income inequality among performers increases because technological change or globalization allows the superstars to serve a much larger market—that is, to scale up production. Specifically, if $w(z)$ denotes the income of an individual of talent z , $p(z)$ denotes the average price for his services, and $q(z)$ is the quantity provided, such that $w(z) = p(z)q(z)$, the standard interpretation of “superstars” is that they have very large markets (a high $q(z)$). This makes such a framework poorly suited for occupations where output is not easily scalable, such as doctors.

In contrast, we focus on such occupations and study an assignment model that is “one-to-one” (or more accurately “a constant-to-one”) where superstars are characterized by a high price $p(z)$ for their services. This makes our paper closer to Gabaix and Landier (2008) who build a “one-to-one” assignment model to study CEOs’ compensation. They argue that since executives’ talent increases the overall productivity of firms, the best CEOs are assigned to the largest firms. They show empirically that the increase in CEO compensation can be

¹Adding network effects, Alder (1985) goes further and writes a model where income can drastically differ among artists of equal talents.

fully attributed to the increase in firms' market size (Grossman (2007) builds a model with similar results). Along the same lines, Määttänen and Terviö (2014) build an assignment model to study house price dispersion and income inequality. They calibrate their model to six US metropolitan areas and find that the increase in inequality has led to an increase in house price dispersion.

Gabaix, Lasry, Lions and Moll (2015) argue that the fast rise in both the share of income held by the top earners and income inequality among these earners requires aggregate shocks to the return of high income earners ("superstar shocks"). Our analysis suggests that even if such shocks only directly affect some occupations they will spill over into other occupations. The original shock may arise from technological change in occupations where span-of-control features are pervasive as suggested by Geerolf (2015).² Globalization can increase the share of income going to the top earners and also increase inequality among these earners (see Manasse and Turini, 2001; Kukharskyy, 2012; Gesbach and Schmutzel, 2014 and Ma, 2015).³

Beyond "superstar" effects, the economics literature has investigated several possible explanations for the rise in top income inequality. Regardless of what the underlying shock or shocks may be, our paper shows how it can spill over into the broader economy. These spillovers could make it difficult to test between the various hypotheses that have been proposed about the key underlying shock. Jones and Kim (2014) and Aghion, Akcigit, Bergeaud, Blundell and Hémous (2015) look at the role played by innovation;⁴ Piketty (2014) argues that top income inequality has increased because of the high returns on capital that a concentrated class of capitalists enjoy; Piketty, Saez and Stantcheva (2014) argue that low marginal income tax rates divert managers' compensation from perks to wages and increase their incentive to bargain for higher wages. Philippon and Reshef (2012) emphasize the role played by the financial sector, and Böhm, Metzger and Strömberg (2015) question whether this premium reflects true talent. For our purposes, all that matters is that consumers have het-

²Geerolf (2015) builds a span of control model to micro-found the fact that firms' size distribution follows Zipf's law. His model naturally leads to "superstar" effects and a bounded distribution of talents can lead to an unbounded distribution of income. Similarly, Garicano and Hubbard (2012) build a span of control model which features positive assortative matching as the most skilled individuals become the most skilled managers who manage large firms which employ the most skilled workers. They use data from the 1992 Census of services on law offices to find support for their model. Yet, span of control issues do not seem *directly* relevant for doctors or real estate agents.

³Yet, none of these papers are able to generate a change in the shape parameter of the Pareto distribution of top incomes through globalization.

⁴Jones and Kim (2014) build a model close to the superstars literature where the distribution of income for top earners is Pareto and results from two forces: the efforts of incumbents to increase their market share and the innovations of entrants who can replace incumbents. Using a panel analysis of US states, Aghion et al. (2015) show empirically that an increase in innovation leads to more top income inequality.

erogeneous preferences, whether or not those preferences reflect realtors’ actual marketing skills or surgeons’ actual cutting skills.

We proceed to present the theoretical model in section 2. Section 3 establishes that higher incomes are associated with higher physician prices. Section 4 introduces our empirical strategy, data, and instrument. Our core empirical results are in section 5. Section 6 examines heterogeneity in spillovers across occupations. In section 7 we use data on physician networks and pricing to directly test the mechanism embedded in the model. We conclude in section 8.

2 Theory

We first present our baseline model. We consider occupations of heterogeneous ability where production is not scalable—there is no mechanism that would allow the more talented to scale up output—and consumption is non-divisible. In this case, we demonstrate that the within-occupation income distribution is tightly linked to that of the general population. To help guide our empirical analysis and determine when we would expect to see spillover effects in the data, we then relax a number of assumptions. “Doctors” will represent occupations where the most skilled workers can produce a good of higher quality but cannot serve more customers than the less skilled, and where customers cannot divide their consumption across several producers. One high-ability doctor is not the same as two decent-ability doctors. Besides doctors, prominent examples are dentists, college professors, and real estate agents.

Although we will refer to the “quality” of the good and the “skill” of the worker, nothing in our model relies on the “high-quality” goods being superior in any objective way. It is merely “quality as perceived by top-earning consumers.” But given the clunkiness of the latter expression, we shorten it to “quality” and refer to the worker’s “skill.”

2.1 The Baseline Model

We consider an economy populated by two types of agents: consumers of mass 1 and (potential) doctors of mass μ_d .

Production. Consumers produce a homogeneous good that serves as the numeraire. They differ in their ability to produce such that a consumer of ability x can produce x units of the homogeneous good. The ability distribution is Pareto such that a consumer is of

ability $X > x$ with probability:

$$P(X > x) = \left(\frac{x_{\min}}{x}\right)^{\alpha_x},$$

with lower bound $x_{\min} = \frac{\alpha_x - 1}{\alpha_x} \hat{x}$ and shape $\alpha_x > 1$, which keeps the mean fixed at \hat{x} when α_x changes. The parameter α_x is an (inverse) measure of the spread of abilities. We will keep α_x exogenous throughout and will capture a general increase in top income inequality by a reduction in α_x . Doctors produce health services and can each serve λ customers, where we impose $\lambda \geq \max(1, \mu_d^{-1})$ so that there are enough doctors to serve everyone. Potential doctors differ in their ability z , according to a Pareto distribution with shape α_z such that they will have ability $Z > z$ with probability:

$$P(Z > z) = \left(\frac{z_{\min}}{z}\right)^{\alpha_z}.$$

All potential doctors can alternatively work as consumer and produce the homogeneous good with ability x_{\min} (see section 2.2.2 for a model where doctors' and consumers' abilities are perfectly correlated). Though the ability of a doctor does not change how many patients she can take care of, it increases the utility benefit that patients get from the health services that are provided.

Consumption. Consumers consume the two goods according to the Cobb-Douglas utility function

$$u(z, c) = z^{\beta_z} c^{1-\beta_z}, \tag{1}$$

where c is the consumption of homogeneous good and z is the quality of the health care (equal to the ability of the doctor providing it).⁵ The notion that medical services are not divisible is captured by the assumption that each consumer needs to consume the services of exactly one doctor. This implies that there will not exist a common price per unit of quality-adjusted medical services. For simplicity, doctors only consume the homogeneous good, an assumption that can easily be generalized (see Appendix B.4.1).

2.1.1 Equilibrium

Consumers. Since a consumer of ability x produces x units of the consumption good, their income must be distributed like their ability. The consumption problem of consumer of

⁵For our purpose, one should think of z as the quality of health care perceived by the consumers at the time when they decide on a doctor. So a pediatrician who can assuage an anxious parent might have a higher z than one with better diagnostic skills but fewer interpersonal skills.

ability x can then be written as:

$$\begin{aligned} \max_{z,c} u(z, c) &= z^{\beta_z} c^{1-\beta_z}, \\ \text{st } \omega(z) + c &\leq x, \end{aligned} \quad (2)$$

where $\omega(z)$ is the price of one unit of medical services by a doctor of ability z .

Taking first order conditions with respect to the quality of the health services consumed and the homogeneous good gives:

$$\omega'(z) z = \frac{\beta_z}{1-\beta_z} [x - \omega(z)]. \quad (3)$$

Since no consumer spends all her income on health care, this equation immediately implies that in equilibrium, $\omega(z)$ must be increasing such that doctors of higher ability earn more per unit of medical services. Importantly, the non-divisibility of medical services implies that doctors are “local monopolists” in that they are in direct competition only with the doctors of slightly higher or lower ability. As a consequence, doctors do not take prices as given, which implies that $\omega(z)$ will in general not be a linear function of z .

As a result, the equilibrium involves positive assortative matching between consumers’ income and doctors’ ability. We denote by $m(z)$ the matching function such that a doctor of ability z will be hired by a consumer whose income is $x = m(z)$. We show that $m(z)$ is an increasing function in Appendix A.1.

Doctors. Since there are (weakly) more doctors than needed the least able doctors will choose to work as consumers. We denote by z_c the ability level of the least able doctor who decides to provide health services so that $m(z)$ is defined over $[z_c, \infty)$ and $m(z_c) = x_{\min}$ (the worst doctor is hired by a consumer with income x_{\min}). Then, market clearing at all quality levels implies that

$$P(X > m(z)) = \lambda \mu_d P(Z > z), \quad \forall z \geq z_c \quad (4)$$

There are $\mu_d P(Z > z)$ doctors with an ability higher than z , each of these doctors can serve λ patients, and there are $P(X > m(z))$ patients whose income is higher than $m(z)$. If $\lambda \mu_d > 1$ then $z_c > z_{\min}$ and if $\lambda \mu_d = 1$ then $z_c = z_{\min}$.

Using the assumption that abilities are Pareto distributed, we can use (4) to obtain the matching function as:

$$m(z) = x_{\min} (\lambda \mu_d)^{-\frac{1}{\alpha_x}} \left(\frac{z}{z_{\min}} \right)^{\frac{\alpha_z}{\alpha_x}}. \quad (5)$$

Intuitively if $\alpha_z > \alpha_x$, so that top talent is “more scarce” among doctors than consumers, then the matching function is convex. This is because it must assign increasingly relatively productive consumers to doctors. At $m(z_c) = x_{\min}$, we obtain the ability of the least able potential doctor working as a doctor: $z_c = (\lambda\mu_d)^{\frac{1}{\alpha_z}} z_{\min}$. This is independent of the consumers’ income distribution because it only depends on quantities.

We denote by $w(z)$ the income of a doctor of ability z and note that $w(z) = \lambda\omega(z)$ since each doctor provides λ units of health services. Furthermore as a potential doctor of ability z_c is indifferent between working as a doctor and in the homogeneous good sector earning a wage equal to x_{\min} , we must have $w(z_c) = x_{\min}$. Now plugging the matching function in (3), we obtain the following differential equation which must be satisfied by the wage function $w(z)$:

$$w'(z)z + \frac{\beta_z}{1-\beta_z}w(z) = \frac{\beta_z}{1-\beta_z}x_{\min} \left(\frac{\lambda^{\alpha_x-1}}{\mu_d} \right)^{\frac{1}{\alpha_x}} \left(\frac{z}{z_{\min}} \right)^{\frac{\alpha_z}{\alpha_x}}. \quad (6)$$

Using the boundary condition at $z = z_c$, we obtain a single solution for the wage profile of doctors. We demonstrate in Appendix A.2 that this function is:

$$w(z) = x_{\min} \left[\frac{\lambda\beta_z\alpha_x}{\alpha_z(1-\beta_z) + \beta_z\alpha_x} \left(\frac{z}{z_c} \right)^{\frac{\alpha_z}{\alpha_x}} + \frac{\alpha_z(1-\beta_z) + \beta_z\alpha_x(1-\lambda)}{\alpha_z(1-\beta_z) + \beta_z\alpha_x} \left(\frac{z_c}{z} \right)^{\frac{\beta_z}{1-\beta_z}} \right]. \quad (7)$$

As expected, the wage profile $w(z)$ is increasing in doctor’s ability z , and $w(z_c) = x_{\min}$. Intuitively, equation (7) consists of two parts: The first term dominates for large $\frac{z}{z_c}$ and ensures an asymptotic Pareto distribution. The second term fulfills the indifference condition for the least able active doctor. Hence, for $\frac{z}{z_c}$ large, we get that

$$w(z) \approx x_{\min} \frac{\lambda\beta_z\alpha_x}{\alpha_z(1-\beta_z) + \beta_z\alpha_x} \left(\frac{z}{z_c} \right)^{\frac{\alpha_z}{\alpha_x}}. \quad (8)$$

Therefore the wage schedule is convex in z if $\alpha_z > \alpha_x$. To understand the intuition, consider again the case where top-talented doctors are scarce ($\alpha_z > \alpha_x$). This implies a fatter tail among consumers than doctors, such that a consumer of twice the income does not have a doctor of twice the ability. Hence, a linear schedule $\omega(z) \propto z$ cannot be an equilibrium as the Cobb-Douglas utility function would require a constant share spent on medical services, which would imply double the payment to a doctor that is not twice as good. For the same reason, the schedule cannot be concave when $\alpha_z > \alpha_x$.

We illustrate this equilibrium in Figure 2. Panel A shows the budget sets and indifference curves for six different consumers, along with the matching function that this equilibrium

generates. For each consumer, the horizontal axis shows consumption c of the homogeneous good, and the vertical axis shows the quality of physician z that the consumer obtains. The dotted curves represent the indifference curves, and the solid curves the budget constraints. With Cobb-Douglas utility, the higher indifference curves are proportionally scaled versions of lower indifference curves—the slopes are constant on any ray out from the origin. But the budget constraints behave very differently: they are curved because there is not a constant price per unit of quality. In this example, additional units of quality have increasing cost. So, for any given budget constraint, the constraint flattens as we move to the left. Income differences lead to parallel shifts left or right in the budget constraint. As a result, the slopes at which the indifference curves are tangent to the budget constraint change for different budget constraints.

Panel B zooms in on the equilibrium region to see these changing slopes more easily. It shows that higher-income consumers are paying higher prices per unit of physician quality at the margin. The skill level of doctors matched to consumers thus increases more slowly for these top consumers: the match function is concave. This is necessary to maintain equilibrium given the higher inequality among consumers. Because there is much more skill dispersion among consumers than among physicians (the distribution of consumers has a fatter right tail), the higher-income consumers have a harder and harder time matching with a highly skilled doctor. So the price per unit quality is increasing, while the match function is concave.

Figure 3 shows two equilibrium outcomes from this matching. Panel A shows the amount that consumers spend on physician care ω , as a function of income x . Panel B shows the wage function that this matching generates for physicians. Because higher-quality doctors are matching with increasingly skilled consumers, and at a faster rate the higher the doctor quality, this wage schedule is convex in z . Combining this wage schedule with the underlying distribution of doctors' ability allows us to characterize income inequality among physicians.

Proposition 1 (*Spillovers*). *Doctors' incomes are asymptotically Pareto distributed with the same shape parameter as the consumers'. In particular an increase in top income inequality for consumers increases top income inequality for doctors.*

To see this result, we first define the relevant distribution. Among the set of practicing doctors (those potential doctors who actually choose to work as doctors), let $P_{doc}(W_d > w_d)$ be the probability that income exceeds w . By definition, this is $P_{doc}(W_d > w_d) = \left(\frac{z_c}{w^{-1}(w_d)}\right)^{\alpha_z}$.

Using equation (8), for w_d large enough, we can approximate this distribution as:

$$P_{doc}(W_d > w_d) \approx \left(\frac{x_{\min} \lambda \beta_z \alpha_x}{\alpha_z (1 - \beta_z) + \beta_z \alpha_x} \frac{1}{w_d} \right)^{\alpha_x}. \quad (9)$$

That is, the income of (actual) doctors is distributed in a Pareto fashion at the top. Importantly, the shape parameter is inherited from the consumers, and is independent of the spread of doctor ability, α_z . Similarly, the income distribution of potential doctors (denoted $P_{pot.doc}$) must then obey (for w_d large enough):

$$P_{pot.doc}(W_d > w_d) \approx \frac{1}{\lambda \mu_d} \left(\frac{x_{\min} \lambda \beta_z \alpha_x}{\alpha_z (1 - \beta_z) + \beta_z \alpha_x} \frac{1}{w_d} \right)^{\alpha_x}.$$

In particular, a decrease in α_x directly translates into a decrease in the Pareto parameter for doctors' income distribution: an increase in inequality among consumers leads to an increase in inequality among doctors. In other words, the increase in top income inequality spills over from one occupation (the consumers) to another (doctors). At the top it also increases the income of doctors—as a decrease in α_x leads to an increase in $P(W_d > w_d)$ for w_d high enough.⁶

Further, a decrease in the mass of potential doctors μ_d (or an increase in the mass of consumers, which we have normalized to 1 here) does not affect inequality among doctors at the top. But it increases the share of doctors who are active (z_c decreases) and their wages (as $w(z)$ increases if z_c decreases).

Taking stock. Proposition 1 establishes the central theoretical result of our paper. For the empirical analysis it is important to establish which assumptions are necessary for the spillover result and which are not. We will do this in subsequent sub-sections. Before pursuing these extensions, however, we consider the implications of our basic result for other key outcomes in this market: health expenditures and welfare inequality.

2.1.2 Implications for spending and welfare

Health expenditures. Health care prices increase sharply at the top. In fact, thanks to the Cobb-Douglas assumption, we obtain that rich consumers spend close to a constant fraction of their income on health. Formally, a consumer with income x spends $\frac{w(m^{-1}(x))}{\lambda}$ on health

⁶Not all doctors benefit, though, as we combine a decrease in α_x with a decrease in x_{\min} to keep the mean constant. As a result the least able active doctor, whose income is x_{\min} , sees a decrease in her income. Had we kept x_{\min} constant so that a decrease in α_x also increases the average consumer income, then all doctors would have gained.

services. Using (5), we obtain that his health spending $h(x)$ must be:

$$h(x) = \frac{\beta_z \alpha_x}{\alpha_z (1 - \beta_z) + \beta_z \alpha_x} x + \frac{1}{\lambda} \frac{\alpha_z (1 - \beta_z) + \beta_z \alpha_x (1 - \lambda)}{\alpha_z (1 - \beta_z) + \beta_z \alpha_x} x_{\min} \left(\frac{x}{x_{\min}} \right)^{-\frac{\alpha_x}{\alpha_z} \frac{\beta_z}{1 - \beta_z}}. \quad (10)$$

Note that health care is a necessity if $\alpha_z (1 - \beta_z) + \beta_z \alpha_x (1 - \lambda) > 0$. This follows from the price gradient consumers face (equation 7), which follows from the prices low-quality and high-quality doctors charge. Low-quality prices are pinned down by the indifference condition for the lowest doctor z_c . High-quality prices are determined purely by the parameters of the utility function and the ability distributions. Specially, consider the case in which $\alpha_z (1 - \beta_z) + \beta_z \alpha_x (1 - \lambda) > 0$: a doctor in the right end of the tail serving a patient of income x earns $\frac{\lambda \beta_z \alpha_x}{\alpha_z (1 - \beta_z) + \beta_z \alpha_x} x$. If the lowest quality doctor were to charge the same share of income, she would earn $\frac{\lambda \beta_z \alpha_x}{\alpha_z (1 - \beta_z) + \beta_z \alpha_x} x_{\min} < x_{\min}$, which would be insufficient to compensate her for her outside option as a consumer earning x_{\min} . Consequently, she must be charging a larger share of patient income, and since everybody consumes the services of exactly one doctor, medical services are a necessity. This is more likely to be the case when the number of patients a doctor can treat, λ , is low or when $\alpha_z > \alpha_x$ so consumers have fatter tails than doctors and doctors charge a smaller part of patient income.

Welfare inequality. The lack of a uniform quality-adjusted price implies that prices vary along the income distribution. Heterogeneity in consumption patterns implies that people at different points of the income distribution face different price indices (Deaton, 1998). Taking this into account implies that a given increase in income inequality translates into a lower increase in welfare inequality. The assignment mechanism implies that as inequality increases, the rich consumers cannot obtain better health services—in fact they pay more for health services of the same quality. This mechanism limits the welfare increase in inequality. Moretti’s (2013) work on real wage inequality across cities can be viewed as proposing a similar assignment mechanism causing high earners to locate in high-cost cities.⁷

To assess this formally, we use a consumption-based measure of welfare: We compute the level of consumption of the homogeneous good $eq(x)$ that, when combined with a fixed level of health quality (namely z_c), gives the same utility to the consumer as what she actually gets in the market. That is, we define $eq(x)$ through $u(z_c, eq(x)) = u(z(x), c(x))$. We then obtain:

⁷Diamond’s (2016) critique argues that the amenities of expensive cities are more valuable to the high earners who choose to live there, so we should not fully adjust incomes for these high costs when calculating welfare. In our context, this critique would apply if high-income consumers had stronger preferences for high-quality doctors than low-income consumers.

Proposition 2 (Welfare inequality). *For x large enough, the welfare measure eq is Pareto-distributed with shape parameter $\alpha_{eq} \equiv \frac{\alpha_x}{1 + \frac{\alpha_x}{\alpha_z} \frac{\beta_z}{1 - \beta_z}}$. Thus $\frac{d \ln \alpha_{eq}}{d \ln \alpha_x} = \frac{1}{1 + \frac{\alpha_x}{\alpha_z} \frac{\beta_z}{1 - \beta_z}}$, implying that an increase in consumers' income inequality translates into a less-than-proportional increase in their welfare inequality. The mitigation is stronger when health services matter more (high β_z) or when doctors' abilities are more unequal (low α_z).*

Ssee Appendix A.3 for the proof.

2.2 Extensions for empirical testing

The baseline model makes a number of important assumptions about the structure of labor markets and production. To devise appropriate empirical tests for spillovers, we need to know which assumptions drive the results and which are innocuous. In the following sub-sections, we consider exactly this question. The formal details for each extension are in Appendix B.

We first establish the necessity of the “non-divisibility” assumption. Section 2.2.1 introduces brewers who produce beer, a divisible good. We show that income inequality of brewers is independent of that of consumers.

Second, in section 2.2.2 we show that the predictions of our model are unchanged if we allow complete mobility across occupations. Specifically, high-ability doctors can work as high-earning consumers.

Third, in preparation for our empirical analysis, we introduce a multi-region model in section 2.2.3. Naturally, without trade or migration between regions top income inequality among doctors must be determined by local consumer income inequality. We show that this remains true even if we allow doctors to move across regions. But if we instead allow cross-region trade of medical services, income inequality for doctors will be the same for all regions. This distinction between “local” services that cannot be traded across regions and tradable “non-local” services will be important for the empirical analysis, which is driven by local variation in general income inequality.

Finally, section 2.2.4 shows the robustness of our core result to other model tweaks. Our results hold if the ability distributions are only asymptotically Pareto distributed, and if doctors consume medical services themselves. We use a more general utility function and show that the spillover effect survives, although the prediction of a spillover elasticity of 1 from Proposition 1 does not generalize.

2.2.1 The Role of the Assortative Matching Mechanism

To highlight the specificity of our mechanism, we add “brewers” to the system. Potential brewers can produce a divisible good, beer. They differ in their ability such that a brewer of ability y can produce quantity y of (quality-adjusted) beer. Their ability distribution is Pareto with shape α_y ; that is a brewer has ability $Y > y$ with probability

$$P(Y > y) = \left(\frac{y_{min}}{y} \right)^{\alpha_y},$$

and α_y is kept constant. If potential brewers do not produce beer they produce x_{min} units of the homogeneous good. We modify the utility function such that $u(z, c, y) = z^{\beta_z} c^{1-\beta_z-\beta_y} y^{\beta_y}$. The first order condition for beer consumption together with a market clearing equation determine the price of beer. As beer is divisible, the beer price p must be taken as given by each producer and brewers’ incomes will simply be given by py . As a result, the income of active beer producers is Pareto distributed with a shape parameter α_y . A change in inequality among consumers can only affect active producers proportionately.⁸ Moreover since beer is divisible, the distribution of the “real” income inequality is unaffected by the presence of beer and the difference between nominal and real income is only driven by the presence of doctors: Proposition 2 still applies and α_{eq} does not change. Consequently, divisibility is essential for spillovers through consumption.

We formalize this discussion in Appendix B.1.

2.2.2 Occupational Mobility

Above we assumed that a potential doctor working as a consumer makes the minimum amount possible as a consumer: x_{min} . In reality it is quite plausible that those succeeding as doctors would have succeeded in other occupations as well (Kirkeboen, Leuven and Mogstad, 2016). To capture this, we now switch to the opposite extreme and assume that there is perfect correlation between abilities as a doctor and as a consumer. We keep the model as before, except we assume that there is a mass 1 of agents who decide whether they want to be doctors or consumers.

We formalize this discussion in Appendix B.2. We show there that, in equilibrium, there

⁸As shown in Appendix B.1, a decrease in α_x increases p for parameters where all potential brewers are actively producing beer. If the extensive margin of brewers is operative the (mean-preserving) increase in income inequality will lower x_{min} and encourage a supply increase of brewers. As a consequence the effect on beer prices, p , from a decrease in α_x is ambiguous.

is always a rank below which some individuals will choose to be doctors. In addition, under reasonable parameter conditions detailed in the appendix, some individuals will also choose to be consumers. So doctor wages grow in proportion to what they could earn as a consumer.

As a result, the model where agents can switch and the one where they cannot are observationally equivalent: doctors’ top income inequality perfectly traces that of the consumers. This is so because even when doctors are not allowed to shift across occupations, the relative reward to the very best doctors adjusts correspondingly with the shift for consumers.

Supply versus demand side effects. In this augmented model, doctors and consumers interact both through a demand effect—consumers are the clients of doctors—and a supply effect—doctors can choose to become consumers. Since the wage level is directly determined by doctors’ outside option, one may think that the mechanism which leads to spillovers in income inequality is very different compared to the demand-side mechanism of the baseline model. But this is not the case. In Appendix B.2.2 we split the role of consumers into two: patients, who only serve the role of consumers of doctor services and an “outside option” which only serves the role of providing doctors with an alternative occupation to providing medical services. We show that the income inequality of doctors is entirely driven by that of their patients and is independent of changes in the income inequality for the outside option. Consequently, the driving force is still the demand side.

2.2.3 Mobility and Open Economy

So far we assumed a closed economy. Since our empirical analysis will rely on local variation in income inequality, we next consider an economy with more than one region. We analyze a case in which medical services can be traded between regions and a case in which doctors can move across regions.

Tradable health care Consider the baseline model of section 2.1. We now assume that there are several regions, $s \in \{1, \dots, S\}$ and we allow some patients (a positive share of consumers in all regions) to purchase their medical services across regions. The distribution of potential doctors’ ability is the same in all regions (and so is the parameter λ). The other parameters, and in particular the Pareto shape parameter of consumers’ income α_x^s is allowed to differ across regions. The cost of health care services must be the same everywhere; otherwise, the consumers who can travel would go to the region with the cheapest health care. Since top talented potential doctors work as doctors (instead of being consumers with income x_{\min}^s), they must all earn the same wage. In all regions, the income distribution of

patients is asymptotically Pareto with parameter $\min_s \alpha_x^s$, because at the very top, overall income inequality follows the income inequality of the most unequal region. (Section B.4.1 elaborates on this logic in more depth.) As a result doctors’ income is asymptotically Pareto with shape parameter $\min_s \alpha_x^s$ in all regions. In other words, income inequality for consumers in the *most unequal* region spills over to doctors in *all* regions.

Empirically, whether the service provided is “local” (non-tradable) or “non-local” (tradable) will depend on the occupations of interest. We will use the results of this section and the previous ones to guide our empirical analysis.

Doctors moving We return again to the baseline model of section 2.1, but we now assume that there are 2 regions, A and B , and that doctors can move across regions.⁹ But medical services are again non-tradable and patients cannot move.¹⁰ The two regions are identical except for the ability distribution of consumers, which is Pareto in both but with possibly different means and shape parameters. Without loss of generality, we assume that $\alpha_x^A < \alpha_x^B$; that is region A is more unequal than region B .

Since the two regions are of equal size, total demand for health services must be the same and on net, no doctors move. On the other hand, most rich patients are in region A , as it is more unequal. As doctors’ incomes increase with the incomes of their patients, nearly all of the most talented doctors will eventually locate in region A . So, in region A , the distribution of doctors’ ability after relocation is asymptotically Pareto. Just as in the baseline model, doctors’ incomes will be asymptotically Pareto distributed with a shape parameter equal to α_x^A .

In region B , free mobility guarantees that doctors of a given quality level earn the same as in region A . However, after the move, the share of doctors that stay in region B decreases with their quality. We obtain:

Proposition 3. *Once doctors have relocated, the income distribution of doctors in region A is asymptotically Pareto with coefficient α_x^A , and the income distribution of doctors in region B is asymptotically Pareto with coefficient α_x^B .*

Consequently, whether doctors can move or not does not alter the *observable* local income distribution, although it does matter considerably for the unobservable local *ability* distribution. Consequently, for our empirical analysis we need not take a stand on whether doctors

⁹The results generalize to more than 2 regions.

¹⁰When doctors are mobile and medical services tradable, the geographic location of agents is undetermined in general, and we would need a full spatial equilibrium model to generate empirical predictions.

are mobile. We cannot empirically distinguish between the free-mobility and no-migration cases using data on income inequality.

We formalize this discussion in Appendix B.3.

2.2.4 Utility Function and Ability Distribution

Appendix B.4 presents the results of two final extensions of the model. First, we consider a change in the ability distribution such that it is only Pareto distributed at the very top. We show that Proposition 1 still applies: a decrease in α_x will directly translate into an increase in top income inequality among doctors.

Second, we generalize the utility function from Cobb-Douglas to CES. In this case, we still obtain that a reduction in α_x leads to a reduction in α_w ; that is, an increase in general top income inequality increases top income inequality among doctors. In this case, however, the magnitude of the spillover changes. Proposition 5 in the appendix shows that, in the most natural case, the Pareto coefficient for wage inequality is $\alpha_w = \frac{\alpha_z}{\left(\frac{\alpha_z}{\alpha_x} - 1\right)^{\frac{1}{\varepsilon} + 1}}$. To see how this responds to local inequality, define $\hat{\alpha} \equiv \frac{d\alpha}{\alpha}$ as the proportional change in the Pareto coefficients, and log-differentiate α_w with respect to α_z and α_x . This yields:

$$\widehat{\alpha}_w = -\frac{\frac{1}{\varepsilon} - 1}{\left(\frac{\alpha_z}{\alpha_x} - 1\right)^{\frac{1}{\varepsilon} + 1}} \widehat{\alpha}_z + \frac{\frac{\alpha_z}{\alpha_x} \frac{1}{\varepsilon}}{\left(\frac{\alpha_z}{\alpha_x} - 1\right)^{\frac{1}{\varepsilon} + 1}} \widehat{\alpha}_x. \quad (11)$$

This tells us how income inequality among doctors responds to percent changes in inequality among consumers ($\widehat{\alpha}_x$) and in physicians' ability z ($\widehat{\alpha}_z$). Consider first the case where $\varepsilon = 1$. This is the limiting case where CES approaches Cobb-Douglas, and equation (11) reduces to $\widehat{\alpha}_w = \widehat{\alpha}_x$. So the spillover elasticity is equal to 1, exactly as in Proposition 1.

Now consider the case in which health care services and the other goods are complements ($\varepsilon < 1$). This is the most natural case empirically, since it implies that the richest consumers spend a smaller share of their income on health, as we demonstrate in section 3.3.¹¹ In this case, the spillover elasticity from consumers' income distribution to the doctors' income distribution is greater than 1. It is also decreasing with the Pareto coefficient of doctors' ability. Mathematically, the coefficient on $\widehat{\alpha}_z$ in equation (11) is negative and that on $\widehat{\alpha}_x$ is greater than 1. This means that a growing spread in doctors' ability (a decrease in α_z) would *reduce* doctors' income inequality (α_w would increase). This is because more top

¹¹This corresponds to a situation where doctors' ability has a fatter tail than consumers' ability. This means that doctors are relatively abundant at the top. With $\varepsilon < 1$, this reduces the income of top doctors, and in fact consumers' expenditure share on health care services tends to 0 for the richest consumers.

doctors would be competing for patients who are spending a declining share of their income on health care, as we move into the tail.

In observational data, α_z and α_x are likely to be positively correlated. Places with more talent dispersion for consumers are likely to also have more talent dispersion for doctors. So if we run an OLS regression of physician income inequality on consumers' income inequality, and we can't control for inequality in doctors' ability (as it's not observable), the coefficient on $\widehat{\alpha}_x$ would be biased downwards. This motivates us to find an instrument for income inequality, which we introduce in section 4.3

2.3 Empirical predictions

To summarize, our model makes the following predictions:

1. High-earning patients are treated by more expensive doctors.
2. An increase in general inequality will lead to an increase in inequality for doctors if they service the general population directly and their services are non-divisible.
3. This is true regardless of whether doctors can move across regions, and regardless of whether doctors' ability is positively correlated with the income they would receive in alternative occupations.
4. If patients can travel easily, doctors' income in each region does not depend on local income inequality.

In the remainder of this paper, we design and execute empirical tests of these predictions. First, section 3 confirms the association between a patient's income and health care spending. We also show that this relationship is driven in part by choosing physicians who charge different prices, rather than consuming different quantities of care. These first results justify the assortative matching that underpins our model.

Second, section 4 introduces our main empirical strategy to test for the spillovers that our model predicts. Section 5 presents these core results.

Section 6 then examines heterogeneity in these spillovers across occupations to determine whether they follow the pattern that our model predicts.

Our final empirical analysis tests the mechanism: do spillovers occur through pricing heterogeneity? In section 7 we introduce data on physician prices and insurance networks. We use these data to illuminate the mechanism in the specific context of physicians.

3 Assortative Matching with Health Spending

Before we introduce the full empirical strategy to identify spillovers, we validate a key element of our model: assortative matching. We use a nationally representative survey and medical claims data to measure the income gradient of medical spending. Our mechanism requires not only that high-income patients spend more on medical care, but also that they visit higher-priced physicians. We therefore measure physician prices, using the method described below, and present the income gradient of physician prices.

To do this, section 3.1 presents institutional background and our method for measuring physician prices. Section 3.2 introduces the data we will use to measure income gradients and physician prices. The results are in section 3.3.

3.1 Institutional background and measurement of physician prices

For multiple reasons, the medical industry in the United States is not perfectly described by the flexible price-setting model of section 2.1. The government plays a substantial role through Medicare and Medicaid, the insurance sector has an important role as an intermediary, there is substantial asymmetric information between patients and doctors, and patients are sometimes willing to travel to seek medical attention. But these features need not substantially impact our analysis. Although the government sets administrative prices for those whose care it pays for directly, providers' negotiations with private insurers generally lead to higher prices (Clemens and Gottlieb, 2017). Even in the presence of asymmetric information, patients often have clear beliefs about who the "best" local doctor in a specific field is (whether or not these beliefs relate to medical skill or health outcomes). And although patients occasionally travel for care, a patient in Dallas is vastly more likely to seek medical care in Dallas than Boston. Furthermore, our empirical strategy more heavily weights large metropolitan areas, which are more likely to have a full portfolio of medical specialties implying less need to travel. To the extent that the medical industry is best described by a national market, the model suggests that this will simply reduce our estimated spillovers.

Despite these complications, the structure of the health insurance industry may embody enough flexibility to incorporate the economic pressures implied by our model. Clemens, Gottlieb and Molnár (2017) show that insurers and physicians frequently negotiate reimbursements as fixed markups over Medicare. They find that, if Medicare sets a reimbursement rate of r_j^M for treatment j , private insurer i 's reimbursement to physician group g for that treatment is generally determined by

$$r_{i,g,j} = \varphi_{i,g} r_j^M.$$

Following their logic, we will use the markups $\varphi_{i,g}$ as a summary measure of the prices charged by physician group g for treating insurer i 's patients. Again following Clemens, Gottlieb and Molnár (2017), we estimate these markups with a regression of the form

$$\ln r_{i,g,j} = \phi_{i,g} + \ln r_j^M + \varepsilon_{i,g,j} \tag{12}$$

on insurance claims data—data that record insurers' payments to provider groups for specific treatments. In equation (12), $\phi_{i,g}$ is an insurer-physician fixed effect which we interpret as the log of the group's markup over Medicare rates. Clemens, Gottlieb and Molnár (2017) show that this regression matches realized physician payments extremely well, and that the levels of these markups reflect economic pressures such as physician market power.¹²

Despite the ability to tailor prices to market conditions, in practice physicians do not always set prices unilaterally as in normal retail markets. The market is not completely decentralized, with each physician setting an individual price to implement the perfectly assortative matching that our model contemplates. Instead, patients purchase insurance and insurers group beneficiaries into different plans, distinguished largely by the breadth of their networks. That is, an expensive Gold plan may have a large network encompassing most physicians in a region, while a cheaper Silver plan may pay physicians lower reimbursements and have a smaller network (Polisky, Cidav and Swanson, 2016).

This network structure provides a mechanism to mediate the heterogeneous consumer preferences that inequality generates. A consumer with a high willingness to pay for physician quality would have to buy an expensive plan, which pays high physician reimbursements. Because of these high reimbursements, many physicians agree to join the plan's network and treat that plan's customers.¹³ A consumer with a lower willingness to pay can buy a cheaper plan, which saves money by paying physicians less—and, as a result, fewer physicians join that network. So the lower-willingness-to-pay patient ends up with less choice of physicians. And the higher-willingness-to-pay patient may end up visiting physicians who receive higher

¹²In the interest of brevity, we refer the reader to Clemens, Gottlieb and Molnár (2017) or Clemens and Gottlieb (2017) for more institutional details about this price setting in the physician context, and Ho (2009) or Gaynor and Town (2011) for the hospital context.

¹³The insurer enforces the network by providing different levels of coverage when patients see in-network and out-of-network providers. Patients who visit an out-of-network physician normally have to pay more, or even all of the cost, out of pocket. In contrast, those who see the in-network physicians that have agreed to accept the network's reimbursement rates generally incur little or no out-of-pocket cost.

reimbursements per visit. We test this below.

3.2 Data on medical spending and physician prices

Medical spending data

We first measure overall health care spending using data from the Medical Expenditure Panel Survey (MEPS). MEPS is a detailed, nationally representative survey of families' health insurance coverage and medical spending. The survey is conducted annually by the Agency for Healthcare Research and Quality and collects information about specific medical expenditures and their costs. We use data from 2008, which includes 31,262 individuals in 12,316 families. We aggregate medical spending to the family level, the level at which income data are also collected. We take logs of both medical spending and family income to measure the income elasticity of spending.

Health insurance claims data

We measure Engel curves for physician *prices* using three source of insurance claims data. The first two sources are the same ones used in Clemens, Gottlieb and Molnár (2017), and are fully described there: Blue Cross/Blue Shield of Texas (BCBS-TX) and the Colorado All-Payer Claims Data (APCD-CO). As a third source, we add All-Payer Claims Data from New Hampshire (APCD-NH). All three datasets have a similar structure: they provide details on patient visits for physician care, and indicate the service provided and the identity of the physician providing treatment (as an actual name or in encrypted form). Crucially, they indicate the amount the physician was paid for each service, the insurer providing coverage, and whether the physician is in-network. They also provide the patient's residential address, down to the zip code level.

Relying on the institutional details described above, we focus on in-network payments. Depending on the details of the patient's insurance contract, and whether the patient has reached an annual deductible or out-of-pocket maximum, the patient or the insurer may have to pay the physician's fee for a particular treatment. But regardless of who is liable, the amount that the physician expects to receive is governed by the rate negotiated between the physician and the insurer. The three databases all provide information on this negotiated amount, known in the industry as the "allowed charge." They indicate the treatment that the fee covers using a 5-digit code established by the Healthcare Common Procedure Coding System (HCPCS).

This provides the information necessary to estimate equation (12). We estimate (12) on each dataset, and then match the physician fixed effects $\hat{\phi}_p$ to the patients who see that physician. We approximate the patients' income using the median family income in their residential zip code.¹⁴ We compute the mean of physician markups among all physician visits from patients in a given zip code z : $\overline{\varphi}_z = \frac{1}{N_{visit}} \sum_{visit \in z} \hat{\phi}_p$. We then regress this on log median family income in the zip code:

$$\overline{\varphi}_z = \mu_0 + \mu_1 \ln(\text{median family income})_z + \varepsilon_z.$$

We run this regression at the zip code level, weighting observations by the number of underlying physician visits in that zip code.

3.3 Results: medical spending, physician prices, and income

Figure 4 Panel A shows the first fact from this analysis: the Engel curve for family medical spending. The graph shows a binned scatterplot, using 20 vigintiles of family income, and the regression line computed on the microdata. The positive relationship is immediately clear, and reflects an elasticity of 0.23. So a 10 percent increase in family income is associated with 2.3 percent more medical spending.

To examine how much of this elasticity reflects differences in prices, as opposed to quantity or composition of care, we move to the medical claims data from APCD-CO. The results are in Panel B of Figure 4. This graph shows an elasticity of 0.19 between physician log markups and log median family income. By comparing this elasticity with that from Panel A, we conclude that 83 percent of the overall spending difference comes from prices.

With these facts in hand, we are comfortable with the model's matching result. We now proceed to test our core prediction: inequality spills over across occupations within a local geographic market.

¹⁴We obtain data on median family income in each Zip Code Tabulation Area (areas that closely approximate zip codes) from IPUMS NHGIS (Manson et al., 2017).

4 Empirical Strategy to Identify Spillovers

4.1 Income data

Our central data set is a combination of the Decennial Census for 1980, 1990 and 2000 and the American Community Survey (ACS) for 2010-2014 (which, combined together, we refer to as 2014) (Ruggles et al. 2015).¹⁵ Going forward, we refer to this combined Census/ACS sample as the “Census data.” We have 5.4 million observations in 1980, growing to 7.4 million observations in 2014, with positive wage income. We use 2010-2014 as opposed to the perhaps more natural 2008-2012 to avoid the immediate aftermath of the Great Recession, which had large impact on top incomes. Data from farther back are a substantially smaller sample so we exclude them from the analysis. We use the 1990 census occupational classification from IPUMS, which consistently assigns occupations throughout the 1980-2014 period. Since we need a reasonable number of observations in order to estimate local inequality, we restrict ourselves to the biggest 253 labor market areas — those with at least 8 observations of physicians in 1980 — for a total of 1,012 observations. We describe additional specific data cleaning choices in Appendix C.

To compute local inequality, we use as our geographic unit the Labor Market Areas (LMAs) defined by Tolbert and Sizer (1996).¹⁶ These are aggregates of the 741 Commuting Zones (CZs) popularized by Dorn (2009). Both commuting zones and labor market areas are defined based on the commuting patterns between counties. But whereas CZs are unrestricted in size, LMAs aggregate CZs to ensure a population of at least 100,000. LMAs are generally sized such that they can be driven through in a matter of a few hours, *e.g.* Los Angeles or New York. Given that our estimation strategy relies on a relatively high number of observations of a particular occupation, labor market areas are a more natural choice.¹⁷

The publicly available income data are censored, generally at around the 99.5th percentile of the overall income distribution, which complicates our estimation of inequality.¹⁸ But,

¹⁵The detailed Decennial Census microdata each comprise 5 per cent of the population, whereas the ACS samples are 1 per cent per year. Combining the years 2010-2014 creates an “artificial” sample of 5 per cent for 2014. The IPUMS extract inflates all numbers to 2014 using the consumer price index.

¹⁶In the raw data, each observation is associated with a particular geographical area (“county groups” for 1980 and “Public Use Microdata Areas” from 1990 onward). These are statistical areas created to ensure confidentiality and have little economic meaning. Alternatively, one could use states, but some local economies, say greater New York City or Washington D.C. span several states. At the same time, some states are too large to meaningfully capture a local economy and others are too small to have sufficient number of observations.

¹⁷Our central results carry through if we instead use Commuting Zones or states as the unit of analysis.

¹⁸Specially, the censoring takes place at \$75,000 for 1980, \$140,000 for 1990, \$175,000 for 2000 and at the

following Armour, Burkhauser and Larrimore (2014), we use the Pareto distribution to work around this problem. Assuming the data are Pareto distributed, the estimated parameter of that distribution provides a measure of inequality.

Specifically, for any random variable \tilde{X} , suppose \tilde{X} follows a Pareto distribution $P(\tilde{X} > \tilde{x}) = \left(\frac{\tilde{x}}{x_{min}}\right)^{-\alpha}$. But the observed value is $x = \min\{\tilde{x}, \bar{x}\}$ for some censoring point, \bar{x} . If we have a sample of draws from this distribution, where N_{cen} are censored and N_{unc} is the set of uncensored observations, we can write the maximum likelihood function of these data as $\mathcal{L}(\alpha) = \left[\prod_{i \in N_{unc}} \alpha \left(\frac{x_{min}}{x_i}\right)^\alpha x_i^{-1} \right] \left(\frac{x_{min}}{\bar{x}}\right)^{\alpha N_{cen}}$. The resulting maximum likelihood estimate for α^{-1} is

$$\frac{1}{\hat{\alpha}} = \frac{1}{N_{unc}} \left[\sum_{i \in N_{unc}} \ln \left(\frac{x_i}{x_{min}} \right) + N_{cen} \ln \left(\frac{\bar{x}}{x_{min}} \right) \right], \quad (13)$$

where N_{unc} is number of uncensored observations. Note, that even without the assumption of a Pareto distribution, equation (13) is a reasonable measure of top income inequality: It is the average log-difference from the minimum possible observation for the uncensored observations, plus the product of the relative number of censored observations times the log distance from the censoring point to the minimum. This will be our measure of income inequality throughout. Armour, Burkhauser and Larrimore (2014) use this method with Current Population Survey data (March supplement) to show that trends in income inequality match those found by Kopczuk, Saez and Song (2010) using uncensored Social Security data.

Part of the value of this approach is that the Pareto parameter can easily be translated into relative incomes at different ranks of the income distribution. For a Pareto distribution with parameter α , the relative income of somebody at the 99th percentile compared to somebody at the 95th percentile is $5^{1/\alpha}$. The Gini coefficient is $(2\alpha - 1)^{-1}$. Guvenen, Karahan, Ozkan, and Song (2015) and Jones and Kim (2014) also employ α^{-1} as a measure of income inequality.

4.2 Summary statistics

Table 1 shows basic descriptive statistics in 2000 for the occupations with the highest presence in the top ten percent of the national income distribution. It reports each occupation's mean income, inequality as measured with the inverse Pareto parameter α^{-1} , and the share of

99.5th percentile at the state level for each individual year 2010-2014. The information provided about the censored variables varies from year to year.

the top 1%, 5%, and 10% that the occupation represents. The income measure here, and throughout the paper, is pre-tax wage and salary income.¹⁹ Table 2 shows the mean, median, 90th, 95th and 98th percentiles among those with positive wage income for each year. All values are in 2014 dollars.²⁰ As discussed in the introduction, the ratios of the 98th to 95th percentiles, and the 95th percentile to the median, have increased during the period. The table also shows the estimate of $\hat{\alpha}^{-1}$ on the top 10 per cent of observations with positive wage income for each year. We present the 98th/95th and 95th/90th ratios implied by the Pareto distribution with the estimated $\hat{\alpha}$ in parentheses. There is a high level of agreement between the predicted and the actual ratios, consistent with a good fit to the Pareto distribution.

Although the censoring point is sufficiently high to allow standard measures of top income inequality to be calculated for most occupations, the high average income of some occupations leads to a larger share being censored. Although the censoring has little impact on the overall distribution, slightly more than 26 per cent of Physicians with positive income are censored in 2000. This implies that we cannot calculate measures of income inequality using high percentiles. But we can still calculate α^{-1} using the assumption of a Pareto distribution. Table 3 shows the result using the top 65 per cent of the uncensored observations.²¹ Consistent with Figure 1, α^{-1} has increased for most occupations in the top during this period. Table D.1 in the Appendix shows the calculated measures of income inequality (using the top 10 per cent of the population) for a number of other occupations, along with the fraction of observations with positive income that are censored. The table shows the same general trend, but with some notable exceptions. In particular there has been little upward trend in top income inequality for truck drivers, sales people, and computer software developers, but substantial increases for financial managers and chief executives. Table D.2 shows which occupations were in the top 1%, 5% and 10% for 1980 and 2014. It is particularly noteworthy that Physicians are increasingly important in the high end of the income distribution over this period. In fact, Physicians are the most common (Census) occupation in the top 1% in

¹⁹The census includes other measures of income, in particular business income, which could be relevant for some occupations. Unfortunately, since wage income and business income are censored separately, estimating a joint distribution for the two would be substantially more complicated. We are in process of getting access to the full uncensored data which will allow us to use total income.

²⁰We use the 98th percentile as the censoring doesn't allow for the calculations of 99th for all years.

²¹Throughout the paper we follow the following rule of thumb when calculating occupation, year, labor market specific measures of income inequality: If there are very few censored observations — say for secretaries — we use the top 10 per cent of the distribution. For occupations that are heavily censored — physicians and dentists — we move the cut-off until we have around twice as many uncensored observations as censored for all labor market areas we use. For Physicians that is the top 65 per cent, for dentists it is top 50 per cent and for Real Estate agents it is top 20 per cent.

2014.

Table 4 shows the size distribution across regions (LMAs) of the number of total observations with positive wage income and physicians with positive wage income. Figure 5 maps the inequality parameters we estimate among physicians and non-physicians for each LMA. Eyeball econometrics suggests a relationship between the two, but we will study this more formally below.

To assess the fit of the Pareto distribution at the LMA-year-occupation level, we use the fact that a Pareto distribution implies a linear relationship between log value and log frequency. Figure 6 shows this relationship for the biggest labor market area for both all occupations (Los Angeles) and physicians specifically (New York). The line shows the predicted number of observations in each bin and the orange dots give the actual number of observations in each bin, both plotted according to the left-side y axis. The right-side y axis gives the corresponding values for the censored values, scaled to ensure that the predicted censored values are on the same line as the uncensored predicted values.²²

Figure 6 Panel A uses the biggest labor market area (Los Angeles) for the year 2000 and bins the income interval between the 90th percentile and the censoring point of \$175,000 into 20 evenly spaced (in logs) bins and plots the predicted log number of observations from the associated Pareto distribution along with the observed log number of observations in each bin (the choice of bins in the figure does not influence any estimation results). The figure further shows the actual and predicted number of censored observations, plotted on the right side axis. The fit for the general population is very close to a straight line and therefore a Pareto distribution. We perform an analogous analysis for the physicians (where New York City is the biggest LMA). Given the lower number of observations overall, and the much higher number of censored observations, we use the top 65 per cent of the positive uncensored observations.²³ The fewer observations implies a fit that is less tight, but there are no systematic deviations from the straight line. Appendix D investigates the quality of this fit in other samples.²⁴

²²Formally, we use the fact that for a dataset with N observations of wages drawn from a Pareto distribution $P(X > x) = \left(\frac{x}{x_{min}}\right)^{-\alpha}$ with a corresponding pdf of $f(x) = \alpha x^{-(\alpha+1)} x_{min}^{\alpha}$, the expected number of observations that have wage income in the interval $[x' - \frac{\Delta}{2}, x' + \frac{\Delta}{2}]$ is $N_{x'} = N \int_{x' - \frac{\Delta}{2}}^{x' + \frac{\Delta}{2}} f(x) dx \simeq N \Delta \alpha x^{-(\alpha+1)} x_{min}^{\alpha}$, giving a negative linear relationship between $\ln N_{x'}$ and $\ln x$. The predicted number of censored observations is $P(X > \bar{x}) = \left(\frac{\bar{x}}{x_{min}}\right)^{-\alpha}$ to which we (arbitrarily) assign the x value $\bar{x} + \frac{\Delta}{2}$ and scale to fit on the same predicted line.

²³Though we carry out the main analysis using the top 65% of observations, Table E.1 in the Appendix shows that the parameter estimate is relatively insensitive to the choice of cut-off.

²⁴Appendix Figures D.1 and D.2 show equivalent graphs for the 20 biggest labor markets in the United

4.3 Empirical strategy

Objective As our main empirical test of the model, we estimate the causal effect of general income inequality on income inequality within a specific occupation. Both in order to test finer model predictions, and in order to have enough empirical variation, we focus on spillovers within a geographic market.

Specifically, we estimate the causal effect of top income inequality in a region s on the top income inequality of a particular subgroup i in region s . Let $\alpha_{o,t,s}^{-1}$ be top income inequality for occupation i at time t for geographical area s and $\alpha_{-o,t,s}^{-1}$ be the corresponding value for the general population in s except for i . Let γ_s be a dummy for the geographical area, γ_t a time dummy, and $X_{t,s}$ a vector of controls, including the area’s population and average income. The regression of interest is then, at the area-occupation-year level:

$$\log(\alpha_{o,t,s}^{-1}) = \gamma_s + \gamma_t + \beta \log(\alpha_{-o,t,s}^{-1}) + X_{t,s}\delta + \epsilon_{o,t,s}. \quad (14)$$

We are centrally interested in β which measures the elasticity of top income inequality for our occupation of interest with respect to the general income inequality.

We will estimate regression (14) by using Census data described above to estimate both $\alpha_{o,t,s}^{-1}$ and $\alpha_{-o,t,s}^{-1}$. This allows us to examine the time period 1980 to 2014 and consider a broad set of occupations.

Instrument One would naturally worry about reverse causality and omitted variables when estimating equation (14). Even controlling for labor market area and year fixed effects, a positive correlation between general income inequality and income inequality for a specific occupation might reflect deregulation, changes in the tax system or common local economic trends—rather than a causal effect from general income inequality to inequality for the occupation of interest. Our mechanism itself could generate reverse causality: physicians’ inequality might spill over into other occupations.

To address these concerns, we use a “shift-share” instrument (following Bartik, 1991) based on the occupational distribution across geographical areas in 1980. We define:

$$I_{-o,t,s} = \sum_{\kappa \in K_{-o}} \omega_{\kappa,1980,s} \alpha_{\kappa,t}^{-1}, \text{ for } t \in \{1980, 1990, 2000, 2014\}, \quad (15)$$

where K_{-o} is the set of the 20 most important occupations in the top ten per cent of the

States.

income distribution nationwide in 1980 (excluding occupation o) and $\omega_{\kappa,1980,s}$ is the share of individuals in occupation κ (as a share of those in any $\kappa \in K_{-o}$) in 1980 in LMA s . In other words, our IV estimation only exploits the changes in labor market income inequality that arises from the occupational distribution in 1980 combined with the nationwide trends in occupational inequality. By using these national trends, our instrument relies on variation associated with national shocks exogenous to the LMA, such as the effects of globalization, technological change or deregulation. This is in line with a decrease in α_x in our theoretical model.

Our primary empirical setup uses two-stage least squares (2SLS) to estimate:

$$\ln(\alpha_{-o,t,s}^{-1}) = \gamma_s + \gamma_t + \pi \ln(I_{-o,t,s}) + X_{t,s}\delta + \epsilon_{o,t,s} \quad (16)$$

$$\ln(\alpha_{o,t,s}^{-1}) = \gamma_s + \gamma_t + \beta \ln(\widehat{\alpha_{-o,t,s}^{-1}}) + X_{t,s}\delta + \epsilon_{o,t,s}. \quad (17)$$

We estimate this in logs to enable more natural comparisons across occupations with different baseline levels of inequality. Using logs also allows for a more direct comparison between the spillover elasticity predicted in the model and the empirical results. The reduced-form equation is:

$$\ln(\alpha_{o,t,s}^{-1}) = \gamma_s + \gamma_t + \pi \ln(I_{-o,t,s}) + X_{t,s}\delta + \epsilon_{o,t,s}. \quad (18)$$

This is identical to the first stage equation (16) except the dependent variable is inequality for doctors instead of consumers.

Throughout our estimation, we cluster standard errors by LMA. Since the inverse of the variance of the MLE estimate for α in equation (13) is proportional to the number of uncensored observations, we weight the regressions by the number of uncensored observations of physicians. Table 5 presents summary statistics for the main regressors.

Occupations We begin our empirical analysis using physicians as the outcome occupation. We consider a number of other occupations, but are limited by the fact that our analysis requires a relatively high number of observations among a high-earning population to measure inequality. We split occupations into two groups: First, we focus on occupations whose output is non-divisible and who primarily operate in local markets: physicians, dentists and real estate agents. Admittedly, some patients do travel for special medical treatment. To

the extent that this creates an integrated market, this would reduce the spillover effects. In the extreme, section 2.2.3 showed that full tradability eliminates local spillovers.

We contrast these with other occupations who have also seen increases in income inequality, but that do not satisfy these conditions: college professors, secretaries, and financial managers. Although professors' output may be non-divisible, they do not operate in a local market—at least not in the right tail of the distribution. Although secretaries operate in local markets, they do not service the general population directly. Finally, consider financial managers. According to the Standard Occupational Classification scheme, financial managers “plan, direct, or coordinate accounting, investing, banking, insurance, securities, and other financial activities of a branch, office, or department of an establishment.” We think of them in two categories: those who manage financial matters for corporations and who are unlikely to be affected by higher local income inequality, and those that manage the financial means of private individuals who are likely to operate in more integrated markets. In either case we expect to see no local spillover effect. Figure 1 Panel B shows the increase in the 99th/90th percentile income ratios for the selected occupations. It demonstrates that the increase in within-occupation top income inequality is a trend outside of the very top of the income distribution as well.

5 Empirical Results: Spillovers

5.1 Testing the model for physicians

Table 6 shows the OLS relationship for physicians. Column 1 estimates equation (14), an OLS regression of physicians' income inequality on general income inequality including year and LMA fixed effects. We find an elasticity of around one-quarter. This estimate remains virtually unchanged in column 2, where we include controls for labor market population and the average wage income among those with positive wage income. Neither control has a significant impact on physicians' income inequality. Column 3 allows the controls to enter more flexibly, by interacting them with year fixed effects. We again see no meaningful change in the main relationship.

Table 7 shows the IV estimates. Columns 1 through 3 show the first stage regression, equation (19). The instrument has a strong predictive power and, along with the time trends, accounts for 82 percent of the variation in the variation for general income inequality (R^2 is computed excluding the LMA fixed effects). The F -statistic for the first stage is in the range of 9 to 11, depending on the controls we use. Columns 2 and 3 add additional controls, just

as in the previous table.

Columns 4 through 6 present the main IV results. The point estimates for the coefficient of interest are 1.00 to 1.29, depending on controls. This is strongly significantly different from 0, and not significantly different from the value of 1 predicted by the simplest model in section 2.1. With our measure of general top income inequality increasing by 27 percent since 1980, and top income inequality for physicians increasing by 31 percent (Tables 2 and 3), an elasticity of 1 suggests that a large share of the rise of income inequality among doctors can be explained by the general increase in income inequality, although the exact fraction is measured with uncertainty. Once again, the controls make little difference to the central estimate.

In section 5.2 we dig into the variation that underpins the Bartik instrument. In section 5.3 we explore the relationship between the OLS and the IV results. In Appendix Table E.2, we show that the results are robust to dramatic changes in sample size. We change the size cut-off between the top 100 LMAs and all LMAs where we are able to estimate physician inequality. The parameter estimate, β , remains significant and between 0.8 and 1.7.

5.2 Understanding the instrument

Traditional vs. flexible Bartik instrument As Beaudry, Green and Sand (2012) and Goldsmith-Pinkham, Sorkin and Swift (2017) show, the shift-share (or Bartik-style) approach relies on the 1980 occupational shares as instruments. The changes in nationwide occupation-level inequality serve as weights. We use this insight to transparently show the source of identifying variation in our setting. To understand this variation, we estimate two first-stage regressions. First, we estimate the traditional Bartik first stage:

$$\alpha_{-o,t,s}^{-1} = \gamma_s + \gamma_t + \pi I_{-o,t,s} + X_{t,s}\delta + \epsilon_{o,t,s}. \quad (19)$$

Second, we run a first stage using the raw occupation shares instead of the Bartik-style summary of those shares. This regression has a separate coefficient on each share, i.e.:

$$\alpha_{-o,t,s}^{-1} = \gamma_s + \gamma_t + \sum_{\kappa \in K_{-o}} \pi_{\kappa,t} \omega_{\kappa,t,s} + X_{t,s}\delta + \epsilon_{o,t,s}. \quad (20)$$

We run these two regressions in levels to ensure that the $\pi_{\kappa,t}$ coefficients in regression

(20) are comparable to the aggregated π coefficient in regression (19). In fact, Goldsmith-Pinkham, Sorkin and Swift (2017) show that the coefficients $\hat{\pi}_{\kappa,t}$ in (20) correspond to the occupation weights in equation (15), up to a constant, multiplied by the coefficient $\hat{\pi}$ from the standard Bartik-style first stage (19). To see this, we will plot the $\hat{\pi}_{\kappa,t}$ coefficients against changes in national inequality for the occupation, $\alpha_{\kappa,t}^{-1}$.

But we first confirm that equations (19) and (20) provide similar instruments by plotting the predicted values $\widehat{\alpha_{-o,t,s}^{-1}}$ from both versions of this estimation against each other. Figure 7 shows that the standard Bartik first stage, equation (19), generates predicted values very similar to the flexible version using individual industry shares, equation (20). Panel A of this figure plots the $\widehat{\alpha_{-o,t,s}^{-1}}$ from equation (19) on the vertical axis, and $\widehat{\alpha_{-o,t,s}^{-1}}$ from equation (20) on the horizontal axis. Panel B shows the residual components of these instruments, *i.e.* disregarding the fixed effects, again from equations (19) and (20) respectively. In both panels we see that the two versions of the first stage yield very similar predicted values. The R^2 of these relationships is 0.94 and 0.64, respectively. So whether we include or exclude the fixed effects, these two sets of predicted values are closely linked.

Sources of variation in the instrument To better understand the variation on which our IV strategy relies, we examine the specific industries that drive the first-stage variation. Since the traditional and flexible versions of the Bartik instrument are so similar, we can learn about the variation underlying the former by looking at which industries make bigger contributions to the flexible version, as estimated by their first stage coefficients $\hat{\pi}_{\kappa,t}$. The first panel of Figure 8 plots the occupation-specific first stage coefficients $\hat{\pi}_{\kappa,t}$ against growth in national inequality for that occupation $\alpha_{\kappa,t}^{-1}$. These first-stage coefficients (plotted along the vertical axis) are nothing more than the naive relationship between *ex ante* occupation shares and subsequent changes in local income inequality. (See section 4.3 above for more details.) The stark positive relationship shown in the figure means that these coefficients are positively associated with what happened in the occupation nationally. This means that there was faster inequality growth in cities with a larger 1980 presence of occupations that subsequently experienced more national growth in income inequality. We estimate a positive relationship with a coefficient close to 1. This buttresses the idea of the Bartik-style instrument: the *ex ante* occupational shares seem to predict growing inequality because of what happens to those occupations nationally.

As a placebo check, we next plot the occupation-specific first stage coefficients $\hat{\pi}_{\kappa,t}$ against another national feature of the occupations: the average top income itself. This is a plau-

sible candidate for an omitted variable that could be correlated with the local occupational distribution and influence physician incomes through a different channel than our spillover. But the second panel of Figure 8 shows no relationship between $\hat{\pi}_{\kappa,t}$ and the growth in the first moment of top income (its mean), as opposed to its second moment (inequality). This provides additional support, beyond the flexible controls included above, for the validity of our instrument.

For further examination of the validity of occupational shares as instruments, we again turn to the suggestions of Goldsmith-Pinkham, Sorkin and Swift (2017). They introduce a test for pre-trends in the Bartik instrument, which we conduct in the appendix.

First stage power The Bartik instrument has strong predictive power: the within-year correlation between $\ln I_{-o,t,s}$ and inequality for non-Physicians, $\ln(\alpha_{-o,t,s}^{-1})$, is between 0.39 and 0.55 for each year (depending on the number of LMAs considered). This correlation is stable across different occupations o , as any one occupation represents a relatively small share of total top income holders.²⁵ When we run the formal first stage, however, the multi-instrument version in (20) has a very weak first stage.²⁶ This may not be surprising, as we have to estimate 57 first-stage coefficients, along with 253 fixed effects, from 1,012 observations. To gain power, we thus use the traditional Bartik instrument, equation (19), as our main first stage. The second stage model in levels is then:

$$\alpha_{o,t,s}^{-1} = \gamma_s + \gamma_t + \beta \widehat{\alpha_{-o,t,s}^{-1}} + X_{t,s} \delta + \epsilon_{o,t,s}. \quad (21)$$

where $\widehat{\alpha_{-o,t,s}^{-1}}$ represents the fitted values from equation (19).

5.3 The relationship between OLS and IV results

Both for physicians and other occupations, the IV results are substantially higher than the OLS correlations. (For example, compare column 4 in Table 7 with column 1 in Table 6.) This likely arises from three sources. First, the augmented model using the CES utility function predicts that the OLS relationship will be biased downwards. We show this mathematically in section 2.2.4; it arises from unobserved correlations between inequality in local doctors' ability and local consumers' ability.

²⁵The qualitative conclusions of our analysis remain unchanged by using a different number of top occupations than 20, although the point estimate of β is somewhat sensitive.

²⁶We have also experimented with using principal components of the occupational share matrix, as suggested by Goldsmith-Pinkham, Sorkin and Swift (2017), but also find weak first stages.

Second, there are numerous potential omitted variables. For example, more unequal places might have higher taxes and spend more public money on health care. This would support incomes of those physicians who are not at the top of the income distribution. Demographics could lead to a downward bias as well, since there may be a negative correlation between rich retirees (whose contribution to physician incomes we ignore) and rich workers (who we measure).

Third, we estimate inequality in the general population with error. Since this is an estimate from a small sample, we would expect the estimated $\alpha_{o,t,s}^{-1}$ to suffer from classical measurement error. This should bias our estimated OLS coefficient downwards.

6 Heterogeneity in Census Results

6.1 Testing the model for occupations with positive predicted spillovers

Moving on from physicians, we analyze two other occupations that are much less regulated and even more local: dentistry and real estate. In Table 8, we study spillovers for dentists just as we did above for physicians. We reach broadly similar conclusions. Again we focus on labor market areas with at least 8 observations in 1980; this severely reduces the number of labor market areas from 253 to 40. Yet we see a pattern broadly similar to that of physicians, albeit with less precision (the p -values are around 0.12). Both OLS and IV point estimates are around twice as high for dentists as for physicians. Though this might reflect the fact that dentistry is more local and prices are less regulated, the point estimates are not significantly distinct and we cannot rule out a difference purely due to sampling error. With a spillover elasticity of 2.8 and a rise in income inequality for dentists that has mirrored that of the general population we substantially over-explain the rise in income inequality for dentists, though with this few observations there is substantial imprecision in the estimate.²⁷

Finally, we use an occupation outside the medical industry: real estate agents. The fee structure in real estate is often proportional to housing prices (Miceli, Pancak and Sirmans, 2007) and the increase in the spread of housing prices is consistent with the increase in income inequality (Määttänen and Terviö, 2014). Real estate is a difficult business to scale up, as

²⁷We also perform the analogous analysis on nurses, for whom top inequality has grown as well (Figure 1 Panel B), though it is less clear that our model would apply to this occupation. Whereas the OLS estimates are similar to physicians, the IV estimate is lower and estimated with more imprecision and we cannot clearly establish spillovers as a cause for the increase in income inequality for nurses. See Table 11 for details.

each house still needs to be shown individually and each transaction negotiated separately. Consequently, one would expect to see spillover effects from general income inequality to real estate agents. Table 9 shows that this is indeed the case. Though the OLS estimates are somewhat lower than for the physicians, the IV estimates are very close. Income inequality for Real Estate agents has increased from 0.45 to 0.69, an increase of 50%. With general income inequality increasing by around 27% the IV estimate suggests that more than half the increase in agents' income inequality can be attributed to the general increase in income inequality.

6.2 Testing the model when spillovers are not predicted

Whereas our theory predicts local spillover effects from general income inequality to the income inequality for occupations such as physicians, dentists and real agents, it predicts no such spillovers for certain other occupations. We perform analogous regressions for financial managers, who, as argued above, do not fit the conditions required for local spillovers. Table 10 shows that this is the case. Though the OLS estimate is positive, the IV estimate is close to zero (and in fact the point estimate is now negative). This also shows that spurious correlation between general inequality and occupational inequality at the local level is likely but that our instrument can address this concern.

Finally, we perform the analysis for two other occupations with substantial increases in top income inequality from Figure 1 Panel B but where our model predicts no spillovers: College Professors and Secretaries.²⁸ Tables 12 and 13 show these results. The top 10% of earners among university professors operate in a national market, and secretaries are not hired by private citizens. For these occupations we find no effect.

Intriguingly, income inequality for secretaries correlates strongly with income inequality for "Chief Executives and Public Administrators" in an OLS regression, shown in Appendix Table E.3. Though we cannot establish causality, this is consistent with a theory analogous to ours, in which CEOs compete for the most skilled secretaries. But at slightly above 10%, the implied elasticity is substantially lower than for other occupations we considered.

²⁸There is a data break in the IPUMS data: For 1980 to 1990 Post-Secondary teachers (those teaching at higher level than high-school) are partly categorized by subject of instruction (code 113-154). From 2000 onward they are not. We collapse all codes 113-154 into 154 for 1980 to 1990.

7 Physician Pricing and Networks

To test the mechanism for spillovers that the model proposes, we rely on the details from section 3.1 of how physician price-setting works in practice in the United States. We exploit data on privately negotiated physician payments and the structure of insurers' networks to directly test the mechanism proposed in our model.

7.1 Data and empirical approach

Insurance network data

In section 3.1, we discussed the idea that patients with higher willingness to pay for their preferred physicians may choose insurance plans with broader, but more expensive, physician networks. Based on the logic of our model, higher income inequality should thus predict more variability in network breadth and physicians' network participation. This provides an institutionally-informed mechanism to transmit income inequality into physician price inequality.

To study inequality in physician networks, we use data collected by the Narrow Networks Project (NNP) at the University of Pennsylvania (Polsky and Weiner, 2015; Zhu, Zhang and Polsky, 2017). This dataset lists the physicians participating in each insurance network for the health insurance exchange plans established under the Affordable Care Act. It reports the physician's identifier, location, and plan participation. We combine it with data on the total number of physicians in each county from the Area Resource File, a standard reference produced by the Department of Health and Human Services.

We construct two primary measures from these datasets. First we consider the share of physicians participating in any exchange plan at all. Since the exchange plans tend to pay lower reimbursements than standard private insurance plans, this is effectively an inverse measure of physicians accepting only high-paying patients. In other words, it measures the uniformity of the health insurance market in a region.

Our second summary is a more direct inequality measure. For each insurance network in a region, we count the number of physicians participating in that network according to the NNP data. We then compute the standard deviation of this measure across networks, which we then normalize by the mean to have a coefficient of variation (CV). This directly measures variability in the size of a network, which provides a mechanism for transmitting heterogeneity in patients' willingness-to-pay into heterogeneity in physician reimbursements.

These measures are only available for a short time horizon: three years, in which the ACA

exchanges were just starting and were constantly in flux. So we only use the most recent year’s data and treat them as a cross-section. We regress the network inequality measures on our standard inequality measure for non-physicians in an LMA, $\log(\alpha_{-o,t,s}^{-1})$.

We standardize all of the network measures, as well as our inequality measure, so regression coefficients can be easily interpreted in terms of standard deviations.

Insurance claims data

To measure inequality in physician prices, we use the three datasets described above in section 3.2. Our estimates of regression (12) generate a distribution of log prices among physicians in a given geographic region. We use these to compute local inequality measures of physician prices. We compute the same $\ln(\alpha^{-1})$ measure we have used throughout the paper, as well as ratios of the 90th to 50th and 75th to 50th percentiles of these markups.

The BCBS-TX data encompass years from 2008-2013, so we will primarily use them as a panel and take differences between the local inequality measures in 2008 and 2013. We regress this short difference on the change in the Bartik instrument from 2000 to 2014, the closest pair of years available for that measure. When using this instrument we continue to run the analysis at the LMA level.

Since we have shorter panels for the APCD datasets, we also run a three-state analysis as a pure cross-section. We amalgamate all of the years of data to form one cross-section, but add richness by computing inequality measures at the finer commuting zone level. We then regress physician price inequality on local income inequality excluding physicians.

7.2 Results

Table 14 presents the network inequality results. The coefficient of 0.72 in column 1 means that variability of network size is 0.74 standard deviations higher in an LMA with 1-standard-deviation higher inequality. Column 2 adds controls for the mix of specialties in an area, and the coefficient falls slightly to 0.62. Both coefficients are statistically significant. Columns 3 and 4 turn to the extensive margin of physician participation in ACA exchange networks. Here, we find that 1-standard-deviation higher inequality is associated with a 0.42-standard-deviation fall in ACA network participation, or 0.23 standard deviations after adding controls. Figure 9 shows these results graphically, using a binned scatterplot.

We report the physician pricing results in Table 15. Columns 1-4 show reduced form regressions of changes in pricing inequality across Texas LMAs from 2008-2013 against changes

in the Bartik instrument. The coefficient of 1.6 in column 1 is statistically indistinguishable from the baseline result for physician income in Table 7. Based on this result, a one-standard-deviation increase in the instrument (0.016) would lead to a one-quarter-standard-deviation increase in pricing inequality growth (0.026/0.112). Column 2 adds controls for the specialty composition in an area, which leads both the coefficient and standard error to approximately double. Results are similar when we use pricing ratios in columns 3 and 4, although column 4 loses statistical significance.

Columns 5 and 6 turn to cross-sectional regressions on a larger sample: three states, and with data at the commuting zone level. These coefficients are not directly comparable to the earlier columns, as the dependent and independent variables are now levels rather than prices. Furthermore, the independent variable is now the realized local income inequality rather than the instrument. These coefficients imply somewhat smaller standardized results: a one-standard deviation increase in inequality (0.115) is associated with one-sixth of a standard deviation higher prices (0.086/0.517) according to column 5. Nevertheless the association remains strongly positive. More unequal areas, and areas with growing predicted inequality, experience more inequality in physician reimbursements.

8 Conclusion

In this paper, we established that an increase in income inequality in one occupation can spill over through consumption to other occupations, such as physicians, dentists and real estate agents, that provide non-divisible services directly to customers. We show that changes in general income inequality at the level of the local labor market area do indeed spill over into these occupations. We distinguish this mechanism by considering other occupations that have seen rises in top income inequality, but that either do not fit our assumptions or operate in a national labor market. Financial managers and college professors experience no spillover effects. This alligns clearly with the predictions of our theory of consumption-driven spillovers. Data on the specific operation of physician markets provides further support for our mechanism.

The magnitude of the key results suggests that this effect may explain most of the increase in income inequality for occupations such as doctors, dentists and real estate agents. As a result, the increase in top income inequality across most occupations observed in the last 40 years may not require a common explanation. Increases in inequality for, say, financial managers or CEOs because of deregulation or globalization may have spilled over to other

high-earning occupations, causing a broader increase in top income inequality.

This analysis has been purely positive, but clearly has normative implications. In particular, Scheuer and Werning (2015) implies that our analysis could be relevant to the study of top income taxation.

References

- Acemoglu, Daron and David Autor**, *Handbook of Labor Economics*, 2011, *40*, 1043–1171.
- Aghion, Philippe, Ufuk Akcigit, Antonin Bergeaud, Richard Blundell, and David Hémous**, “Innovation and Top Income Inequality,” NBER Working Paper No. 21247, National Bureau of Economic Research 2015. NBER wp 21247.
- Alder, Moshe**, “Stardom and Talent,” *The American Economic Review*, 1985, *75(1)*, 208–212.
- Armour, Philip, Richard Burkhauser, and Jeff Larrimore**, *Economic Inquiry*, 2016, *54*, 1263–1273.
- Bakija, Jon, Adam Cole, and Bradley Heim**, “Jobs and Income Growth of Top Earners and the Causes of Changing Income Inequality: Evidence from U.S. Tax Return Data,” 2010.
- Bartik, Timothy J.**, “Boon or Boondoggle? The Debate Over State and Local Economic Development Policies,” *Upjohn Institute for Employment Research*, 1991, pp. 1–16.
- Beaudry, Paul, David A Green, and Benjamin Sand**, “Does industrial composition matter for wages? a test of search and bargaining theory,” *Econometrica*, 2012, *80 (3)*, 1063–1104.
- Böhm, Michael, Daniel Metzger, and Per Strömberg**, ““Since you’re so rich, you must be really smart”: Talent and the Finance Wage Premium,” Working Paper 313, Swedish Riksbank November 2015.
- Clemens, Jeffrey and Joshua D. Gottlieb**, “In the Shadow of a Giant: Medicare’s Influence on Private Payment Systems,” *Journal of Political Economy*, February 2017, *125 (1)*, 1–39.
- , – , and **Tímea Laura Molnár**, “Do Health Insurers Innovate? Evidence From the Anatomy of Physician Payments,” *Journal of Health Economics*, September 2017, *55*, 153–167.
- Deaton, Angus**, “Getting Prices Right: What Should be Done?,” 1998, *12 (1)*, 37–46.

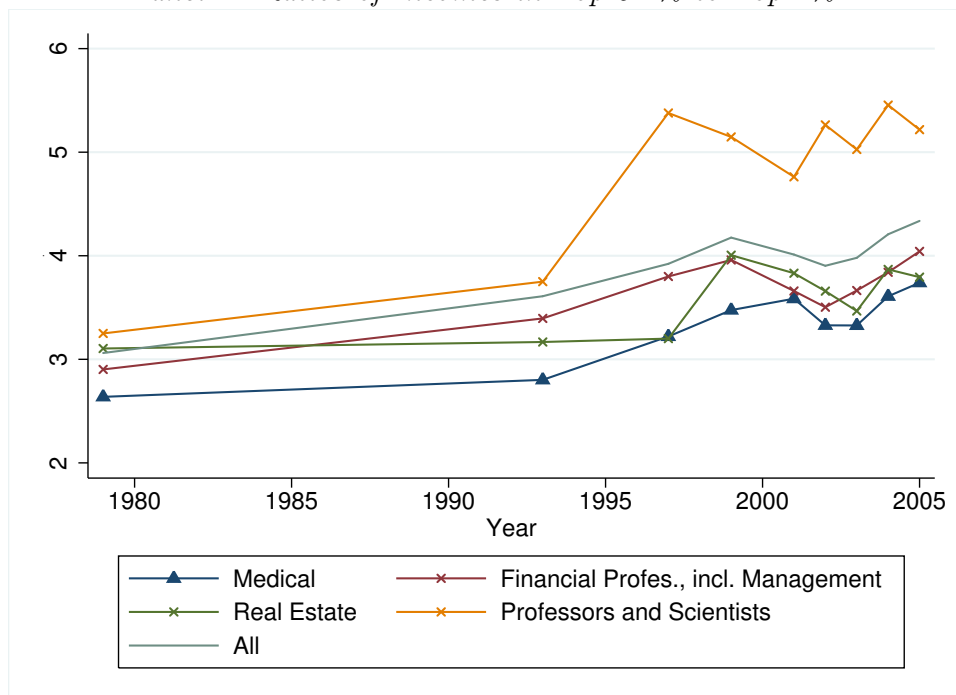
- Diamond, Rebecca**, “The determinants and welfare implications of US workers’ diverging location choices by skill: 1980–2000,” *American Economic Review*, 2016, *106* (3), 479–524.
- Gabaix, Xavier and Augustin Landier**, “Why has CEO pay increased so much?,” *The Quarterly Journal of Economics*, 2008, *123* (1), 49–100.
- , **Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll**, “The Dynamics of Inequality,” 2015. NBER wp 21363.
- Garicano, Luis and Thomas Hubbard**, “Learning about the nature of production from equilibrium assignment patterns,” *Journal of Economic Behavior & Organization*, 2012, *84*, 136–153.
- Gaynor, Martin and Robert J. Town**, “Provider Competition,” in Thomas G. McGuire, Mark Pauly, and Pedro P. Barros, eds., *Handbook of Health Economics, Vol 2*, Amsterdam: Elsevier, 2011.
- Geerolf, Francois**, “A Static and Microfounded Theory of Zipf’s Law for Firms and of the Top Labor Income Distribution,” 2014.
- Gesbach, Hans and Armin Schmutzler**, “Does globalization create superstars? A simple theory of managerial wages,” *European Economic Review*, 2014, *71*, 34–51.
- Goldin, Claudia and Lawrence Katz**, *The Race between Education and Technology* 2010.
- Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift**, “Bartik Instruments: What, When, Why and How,” July 2017. Available online at http://paulgp.github.io/papers/bartik_gpss.pdf (accessed August 28, 2017).
- Grossman, Volker**, “Firm size, productivity and manager wages: a job assignment approach,” *The B.E. Journal of Theoretical Economics*, 2007, *7* (1).
- Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song**, “What do data on millions of U.S. workers reveal about life-cycle earnings risk,” 2015. NBER working paper 20913.
- Ho, Katherine**, “Insurer-provider networks in the medical care market,” *The American Economic Review*, 2009, *99* (1), 393–430.

- Jones, Charles and Jihee Kim**, “A Schumpeterian model of Top income inequality,” 2014. NBER working paper 20637.
- Kaplan, Steven and Joshua Rauh**, “It’s the market: the broad-based rise in the return to top talent,” *Journal of Economic Perspectives*, 2013, 27 (3), 35–56.
- Kirkeboen, Lars J, Edwin Leuven, and Magne Mogstad**, “Field of study, earnings, and self-selection,” *The Quarterly Journal of Economics*, 2016, 131 (3), 1057–1111.
- Kopczuk, Wojciech, Emmanuel Saez, and Jae Song**, *Quarterly Journal of Economics*, 2010, 125 (1), 91–128.
- Kukharsky, Bohdan**, “Trade, Superstars, and Welfare,” 2012. BGPE Discussion Paper No 120.
- Ma, Lin**, “Globalization and Top Income Shares,” 2015.
- Määttänen, Niku and Marko Terviö**, “Income distribution and housing prices: an assignment model approach,” *Journal of Economic Theory*, 2014, 151, 381–410.
- Manasse, Paolo and Alessandro Turrini**, *Journal of International Economics*, 2001, 54, 97–117.
- Manson, Steven, Jonathan Schroeder, David Van Riper, and Steven Ruggles**, *IPUMS National Historical Geographic Information System: Version 12.0*, Minneapolis: University of Minnesota, 2017.
- Miceli, Thomas, Katherine Pancak, and C Sirmans**, “Is the Compensation Model for Real Estate Brokers Obsolete?,” *Journal of Real Estate Finance and Economics*, 2007, 35 (7).
- Moretti, Enrico**, “Real Wage Inequality,” *American Economic Journal: Applied Economics*, 2013, 5 (1), 65–103.
- Pareto, Vilfredo**, *Cours d’Economie Politique* 1896.
- Philippon, Thomas and Arief Reshef**, “Wages and Human Capital in the U.S. Finance Industry: 1909-2006,” *Quarterly Journal of Economics*, 2012, 127 (4), 1551–1609.
- Piketty, Thomas**, *Capital in the Twenty-First Century*, Harvard University Press, 2014.

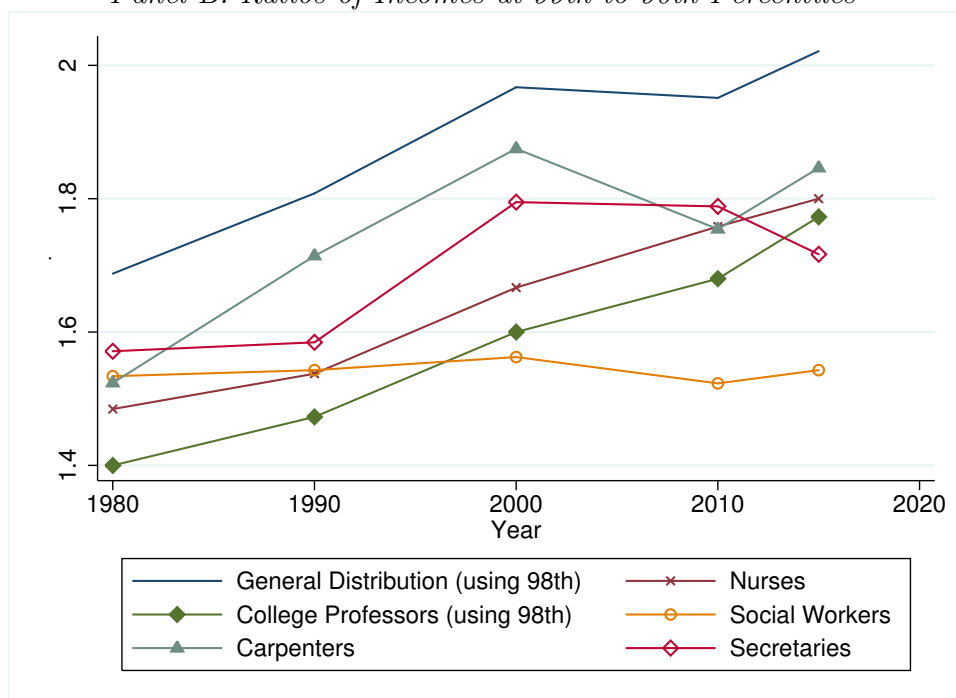
- **and Emmanuel Saez**, “Income Inequality in the United States: 1913-1998 , 1, 1-39,” *Quarterly Journal of Economics*, 2003, *118* (1), 1–41.
 - , **Anthony B. Atkinson, and Emmanuel Saez**, “Top Income in the Long Run of History,” *Journal of Economic Literature*, 2011, *49* (1), 3–71.
 - , **Emmanuel Saez, and Stefanie Stantcheva**, “Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities,” *American Economic Journal: Economic Policy*, 2014, *6*, 230–271.
- Polsky, Daniel and Janet Weiner**, “The skinny on narrow networks in health insurance marketplace plans,” Data Briefs, Leonard Davis Institute June 2015. Available online at <https://ldi.upenn.edu/sites/default/files/pdf/the-skinny-on-narrow-networks.pdf>.
- , **Zuleyha Cidav, and Ashley Swanson**, “Marketplace Plans With Narrow Physician Networks Feature Lower Monthly Premiums Than Plans With Larger Networks,” *Health Affairs*, 2016, *35* (10), 1842–1848.
- Rosen, S.**, “The Economics of Superstars,” *American Economic Review*, 1981, *71*, 485–858.
- Ruggles, Steve, Ronald Goeken, Josiah Grover, and Matthew Sobek**, *Integrated Public Use Microdata Series: Version 6.0*, Minneapolis: University of Minnesota, 2015.
- Sattinger, Michael**, “Assignment models of the distribution of earnings,” *Journal of Economic Literature*, 1993, *31* (2), 831–880.
- Scheuer, Florian and Ivan Werning**, “The Taxation of Superstars,” 2015. NBER wp 21323.
- Zhu, Jane M, Yuehan Zhang, and Daniel Polsky**, “Networks In ACA Marketplaces Are Narrower For Mental Health Care Than For Primary Care,” *Health Affairs*, 2017, *36* (9), 1624–1631.

Figure 1: Relative Income for Selected Occupations Over Time

Panel A: Ratios of Incomes in Top 0.1% to Top 1%

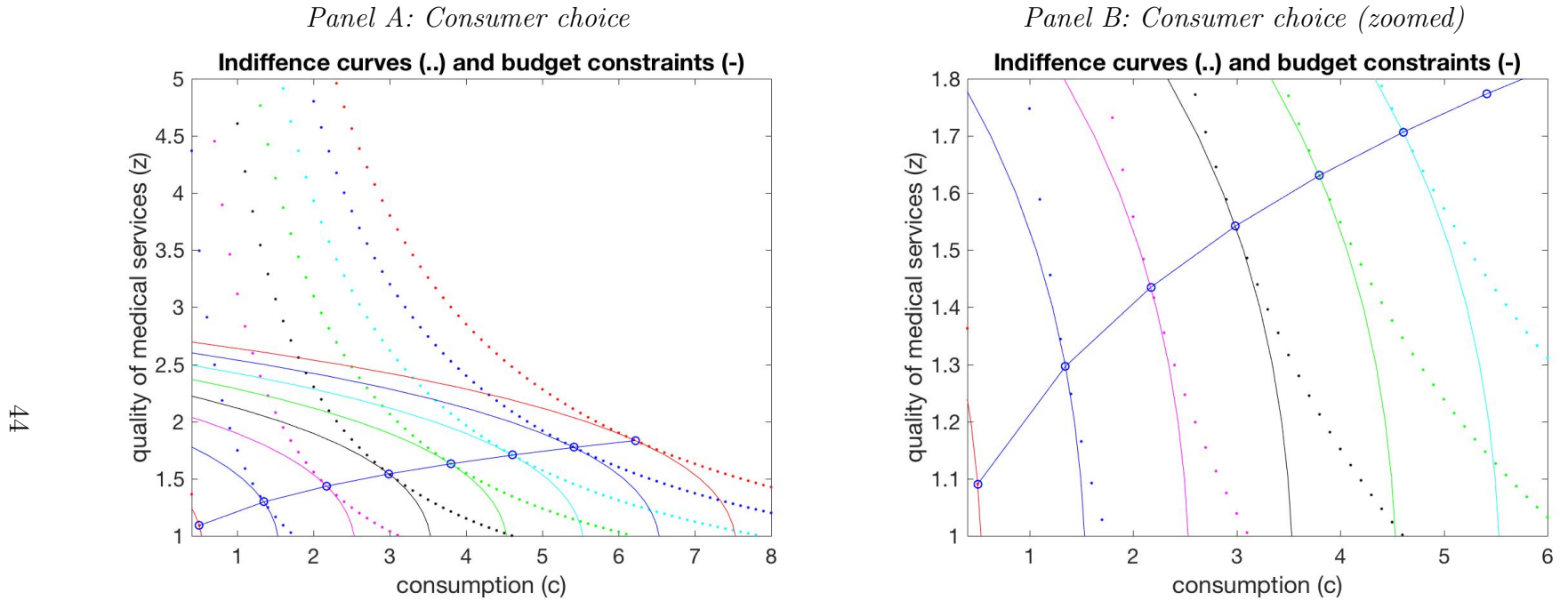


Panel B: Ratios of Incomes at 99th to 90th Percentiles



Notes: Panel A shows the ratio of mean earnings among those in the top 0.1% of the income distribution relative to those in the top 1%, for selected occupations. “All” refers to the full income distribution, not just the occupations shown here. Source: Bakija, Cole and Heim (2012). Panel B shows the ratio of the 99th percentile of the income distribution to the 90th percentile, for selected occupations. The sample consists of employed workers with positive wage income. Censoring prevents the calculation of the 99th percentile for the general distribution as well as for college professors, so we show the 98th percentile instead for those samples. Source: Authors’ calculations using Decennial Census and American Community Survey data from IPUMS (Ruggles et al. 2015).

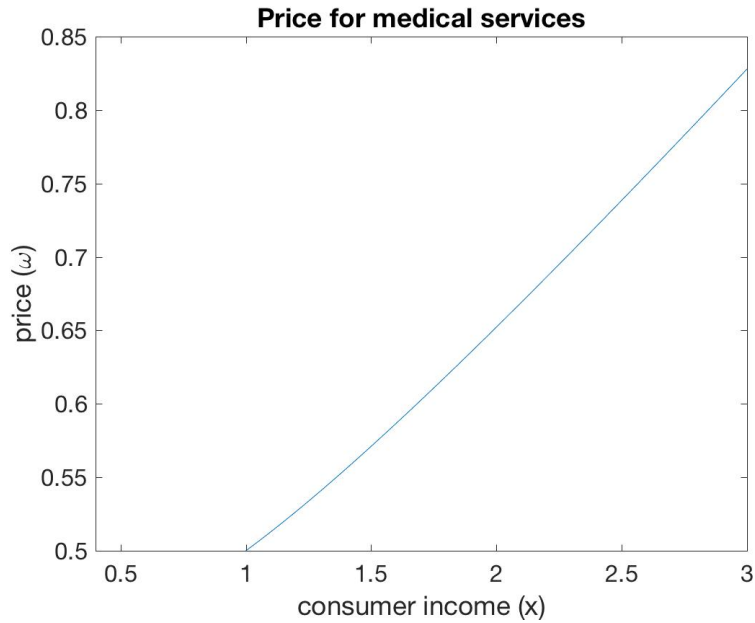
Figure 2: Equilibrium in the Model



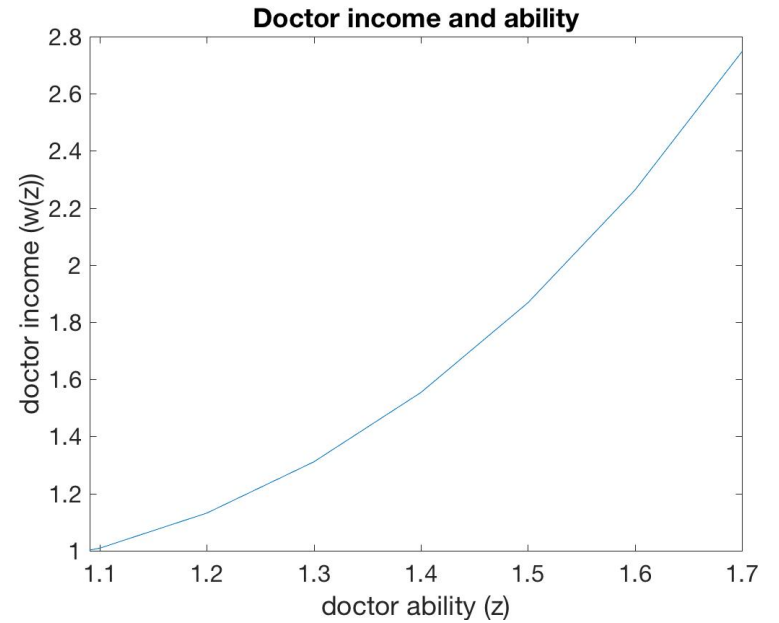
Notes: This figure illustrates the matching mechanism in the model. We compute the solution to our model with the following parameters: $\alpha_z = 8$, $\alpha_x = 2$, $\beta_z = \frac{1}{2}$, $\lambda = 2$, $\mu_d = 1$. The lowest-ability active doctor is approximately $z_c \approx 1.0905$. Panel A shows the budget sets and indifference curves for six different consumers, along with the matching function that this equilibrium generates. The horizontal axis shows consumption c of the homogeneous good, and the vertical axis shows the quality of physician z that the consumer obtains. The dotted curves represent the indifference curves, and the solid curves the budget constraints. The budget constraints are curved because there is not a constant price per unit of quality; in this example, additional units of quality have increasing cost. So, for any given budget constraint, the constraint flattens as we move to the left. Because there is much more skill dispersion among consumers than among physicians (the distribution of consumers has a fatter right tail), the higher-income consumers have a harder and harder time matching with a highly skilled doctor. So the matching function is concave. Panel B zooms in to see this concavity more easily. Panel C shows the wage function that this matching generates for physicians, which is convex in z .

Figure 3: Incomes and Spending in Equilibrium

Panel A: Consumer income and medical spending



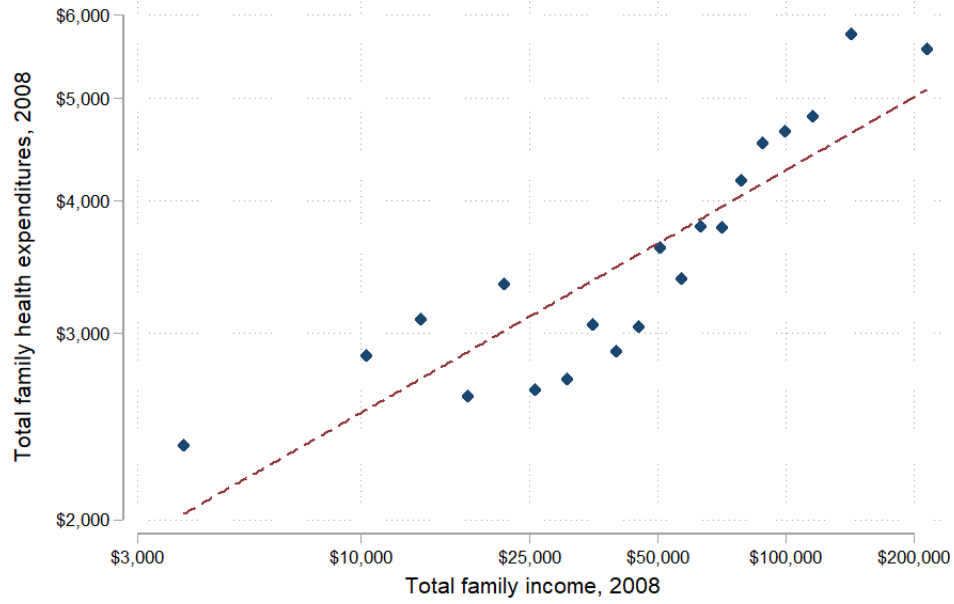
Panel B: Doctor income and ability



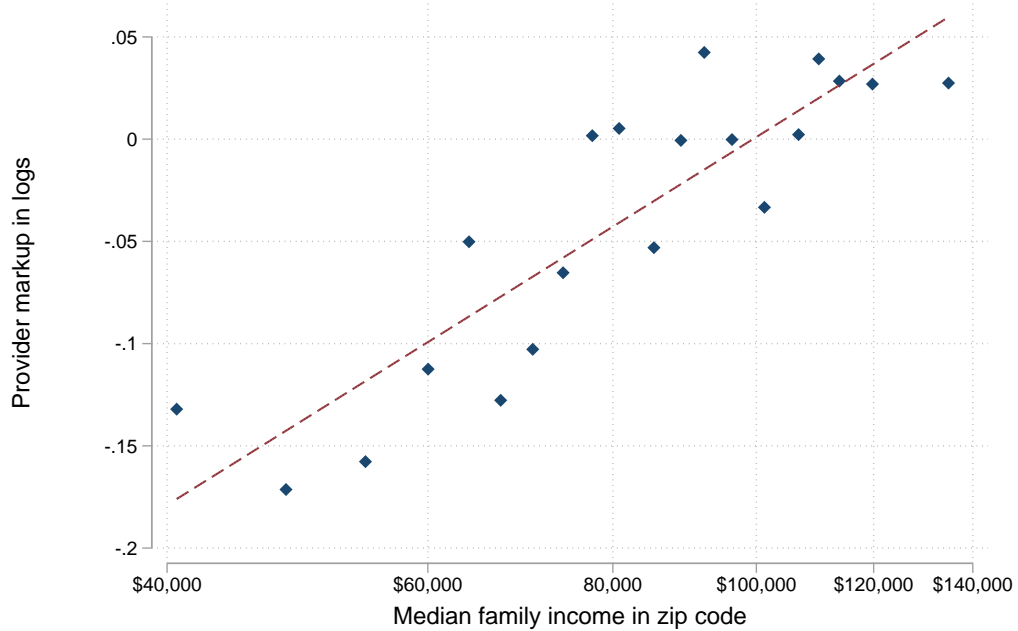
Notes: This figure illustrates some of the outcomes of the model. We compute the solution to our model with the following parameters: $\alpha_z = 8$, $\alpha_x = 2$, $\beta_z = \frac{1}{2}$, $\lambda = 2$, $\mu_d = 1$. The lowest-ability active doctor is approximately $z_c \approx 1.0905$. Panel A shows the price ω that consumer of income x pays for physician care. This is convex in income, as we saw from the slopes of the indifference curves in Figure 2. Panel B shows the wage function $w(z)$ that determines income for a doctor of ability z . This is also convex in ability, and in fact the physician income distribution inherits the Pareto parameter of the consumers' income distribution (asymptotically).

Figure 4: Engel Curves for Medical Spending and Physician Prices

Panel A: Health Expenditures and Family Income



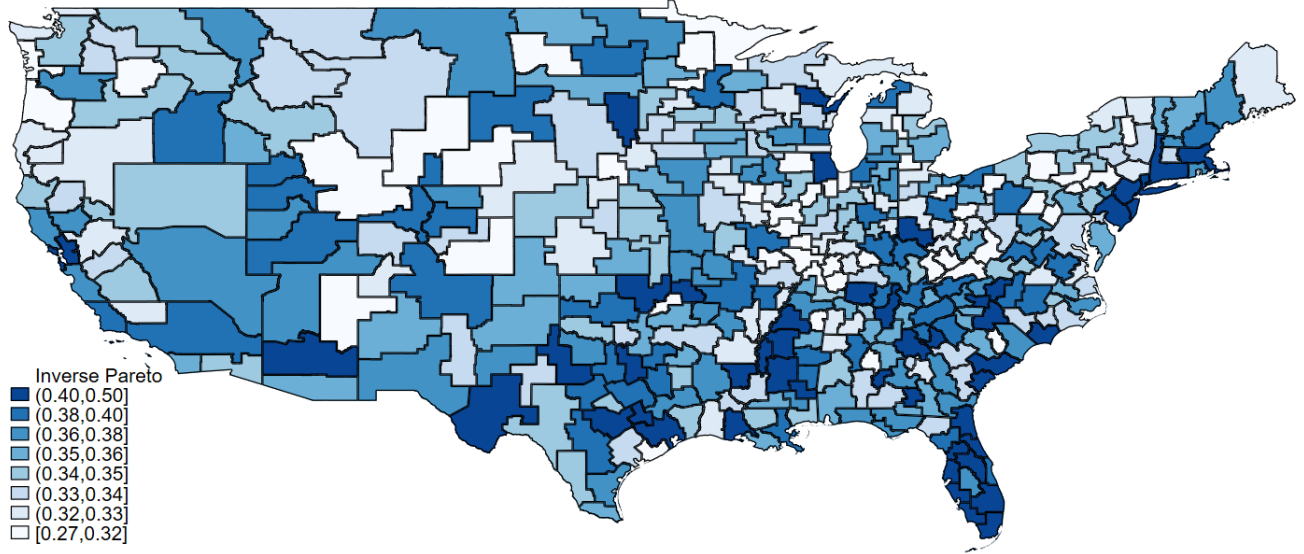
Panel B: Physician Prices and Median Family Income



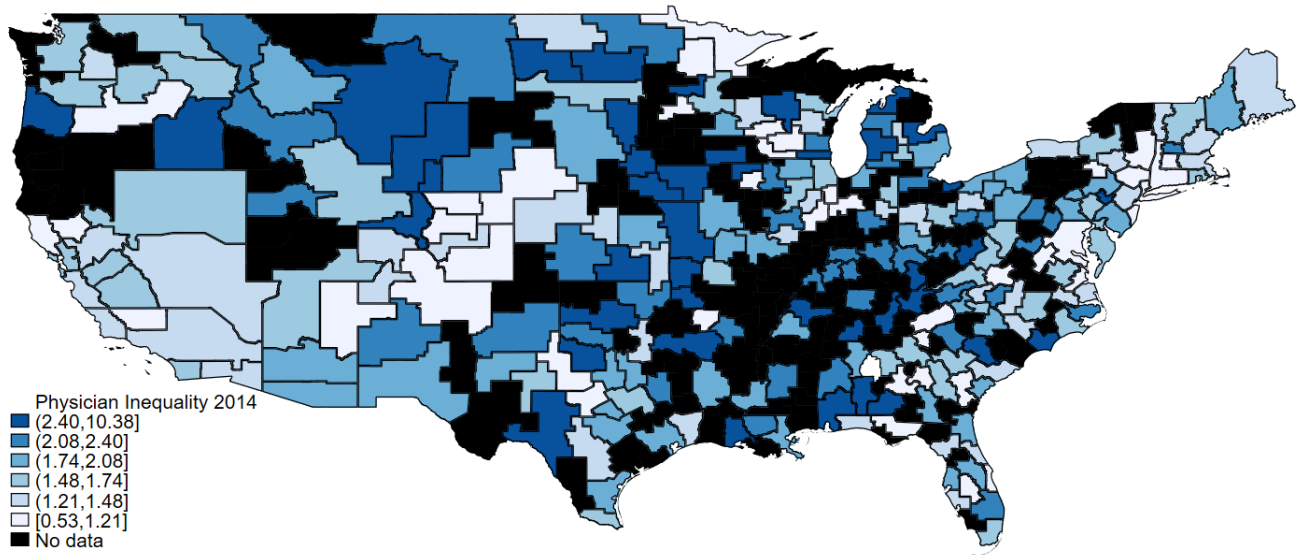
Notes: Panel A shows the relationship between annual medical spending and annual income, both in logs. (The dollar values shown are computed by exponentiating the actual value along each axis). The circles show mean values for each bin, computed as twenty vigintiles of family income. The dashed line is a linear regression estimate on the micro-data, and reflects an elasticity of 0.23 (so a 10% increase in income is associated with 2.3% extra medical spending). Panel B shows an analogous relationship between log markups charged by the physician who treats a patient, and family income (as proxied by the income in the patient's zip code of residence). This panel also shows a binned scatterplot with 20 vigintiles, and the regression line reflects an elasticity of 0.19. Source: Authors' calculations using data from the Medical Expenditure Panel Survey and Colorado All-Payer Claims Data.

Figure 5: Inequality Across the United States

Panel A: Income Inequality Among the General Population

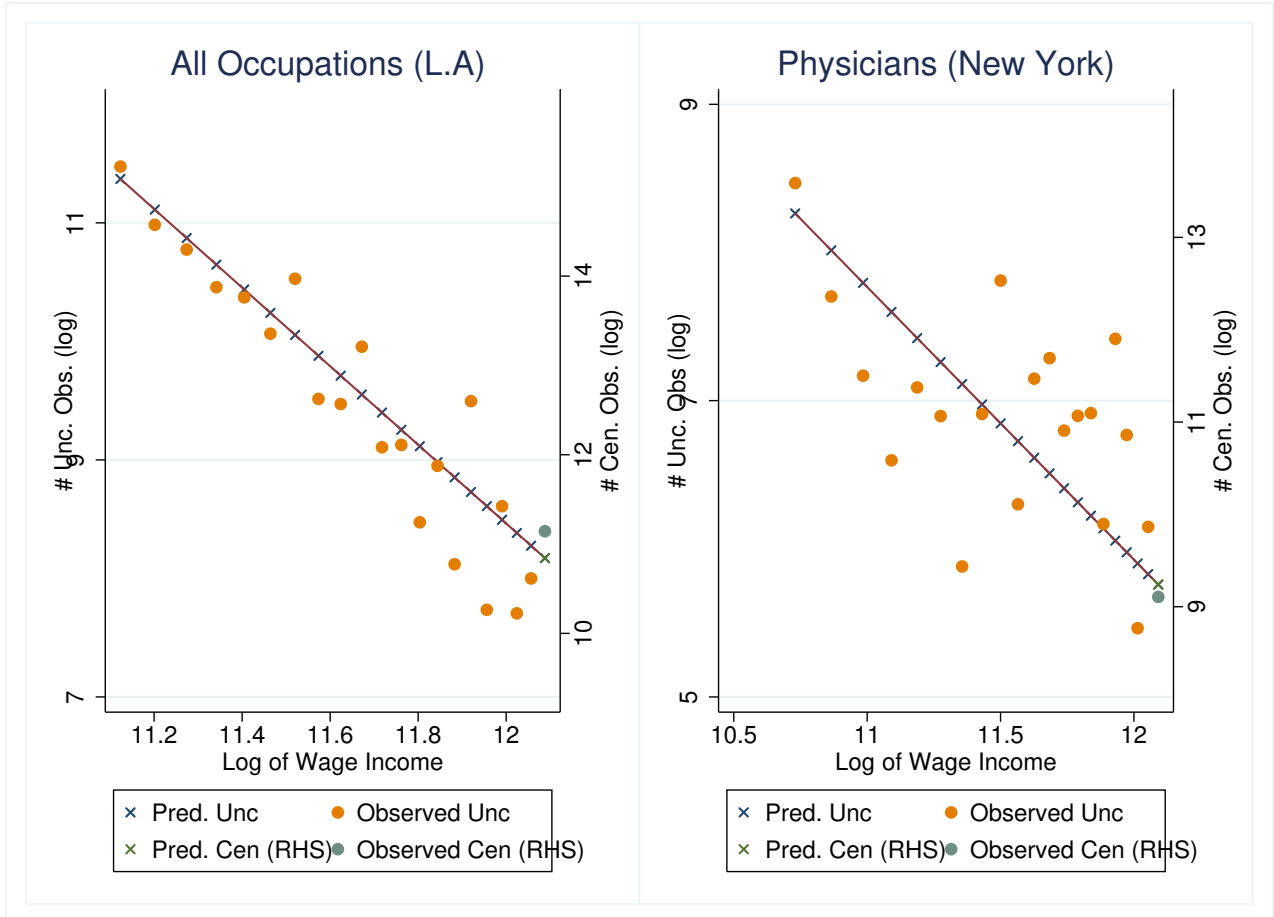


Panel B: Income Inequality Among Physicians



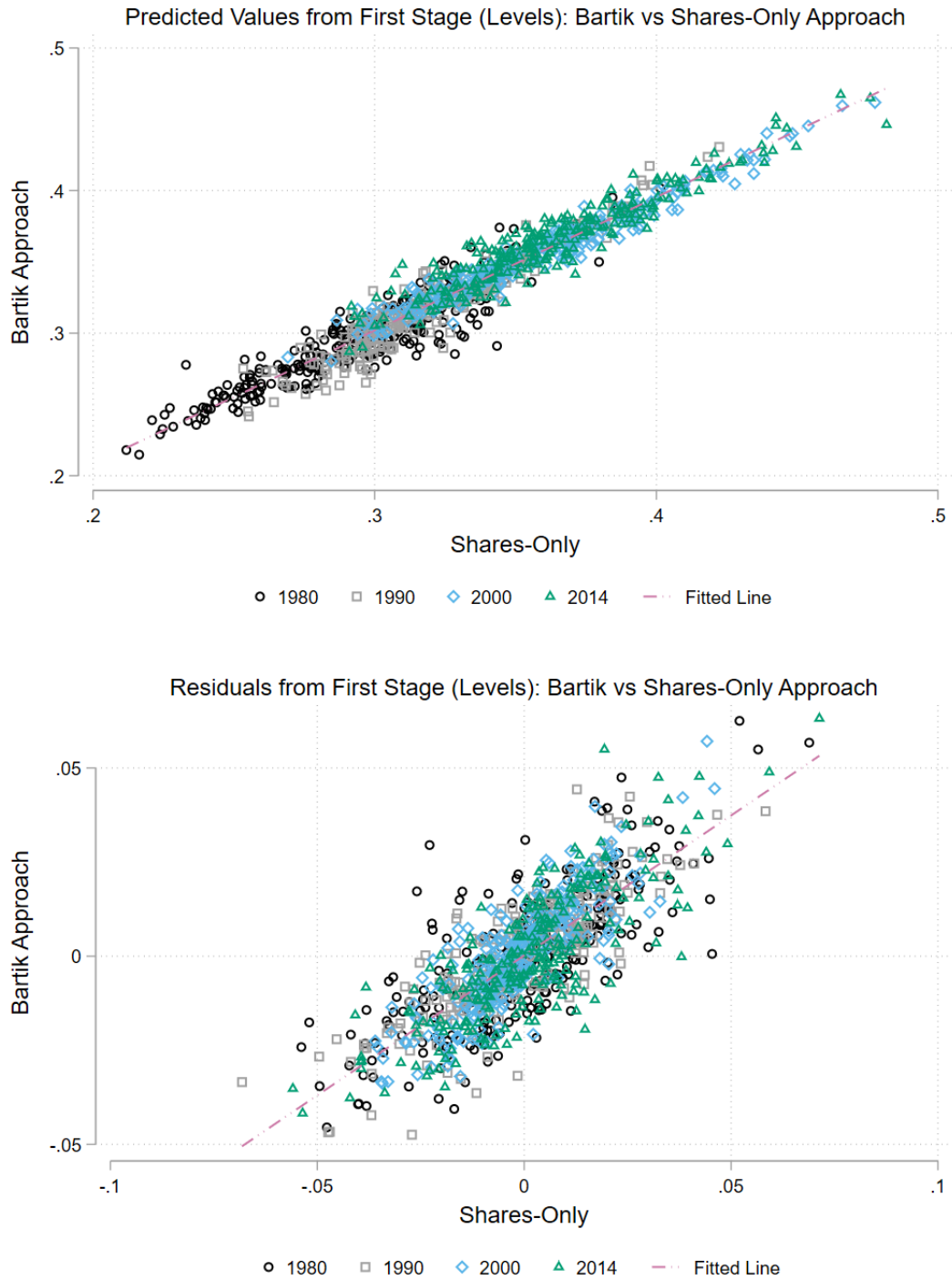
Notes: This figure shows our measure of income inequality by labor market area in American Community Survey data from 2010-2014. We estimate the Pareto parameter of the local income distribution (in Panel A), or the local income distribution for physicians (Panel B), and plot the 95/90th percentile ratio predicted by that Pareto parameter. Source: Authors' calculations using American Community Survey data from IPUMS (Ruggles et al. 2015).

Figure 6: Fit of the Pareto Distribution



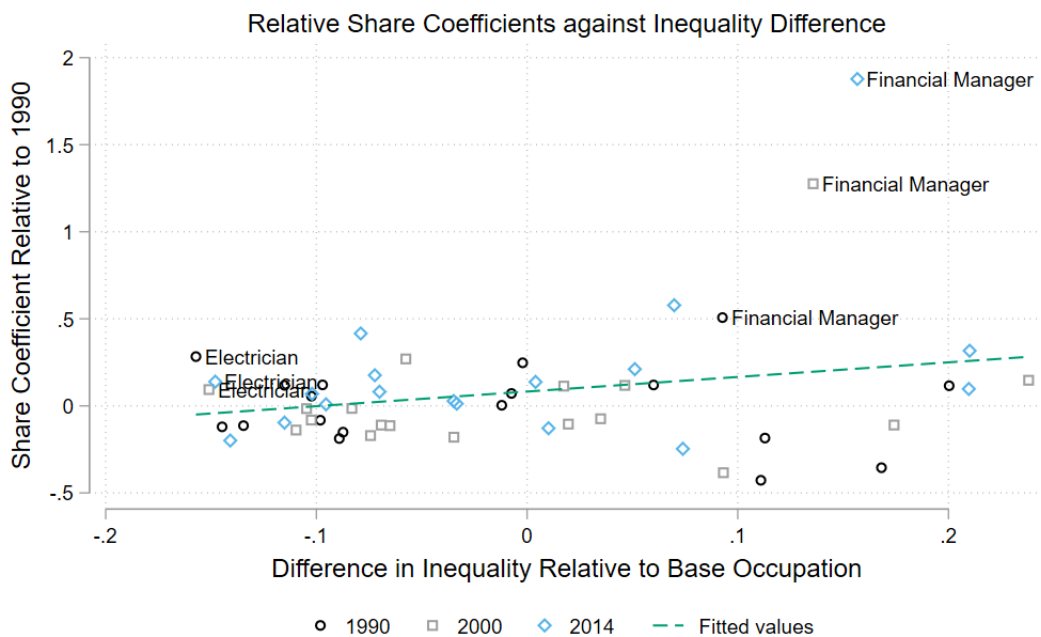
Notes: This figure shows the quality of fit of the empirical income distribution to the Pareto distribution for two samples in data from 2000. The left panel shows the full sample in Los Angeles, and the right panel shows physicians in New York. The horizontal axis shows the log incomes, which are binned, and the vertical axis shows the log number of observations. The lines show the predicted distribution if it were entirely Pareto, and the dots show the empirical sample sizes. The color changes for the highest income value shown, which includes the observations where income is censored, and which is plotted relative to the axis on the right. Source: Authors' calculations using Decennial Census data from IPUMS (Ruggles et al. 2015).

Figure 7: Comparing the Traditional and Flexible Bartik Instruments

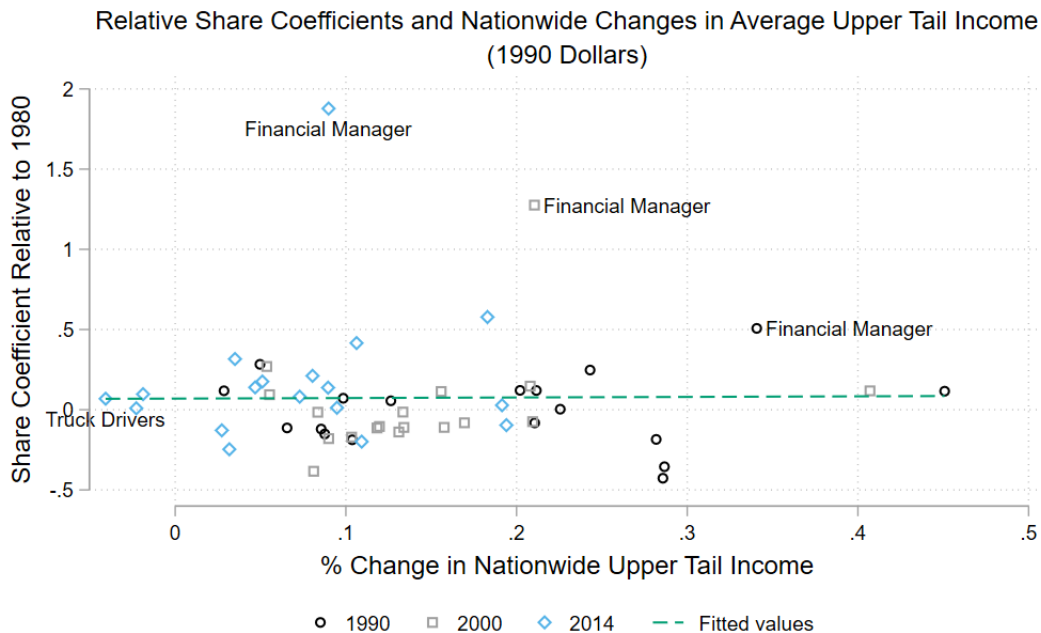


Notes: This figure shows the similarity in predicted first stage values from the two versions of the Bartik instrument. In both panels, the horizontal axis shows a component of the predicted value $\widehat{\alpha_{-o,t,s}^{-1}}$ from the flexible version of the Bartik instrument (shown in equation [20]), while the vertical axis shows the same from the traditional Bartik instrument (shown in equation [19]). Panel A shows the full predicted values, while Panel B shows the part that is not explained by fixed effects. The R^2 of these relationships are 0.94 and 0.64.

Figure 8: Visualizing the Bartik Instrument



The slope is 0.84 with a standard error of 0.44. The coefficient on the bartik IV is .7.



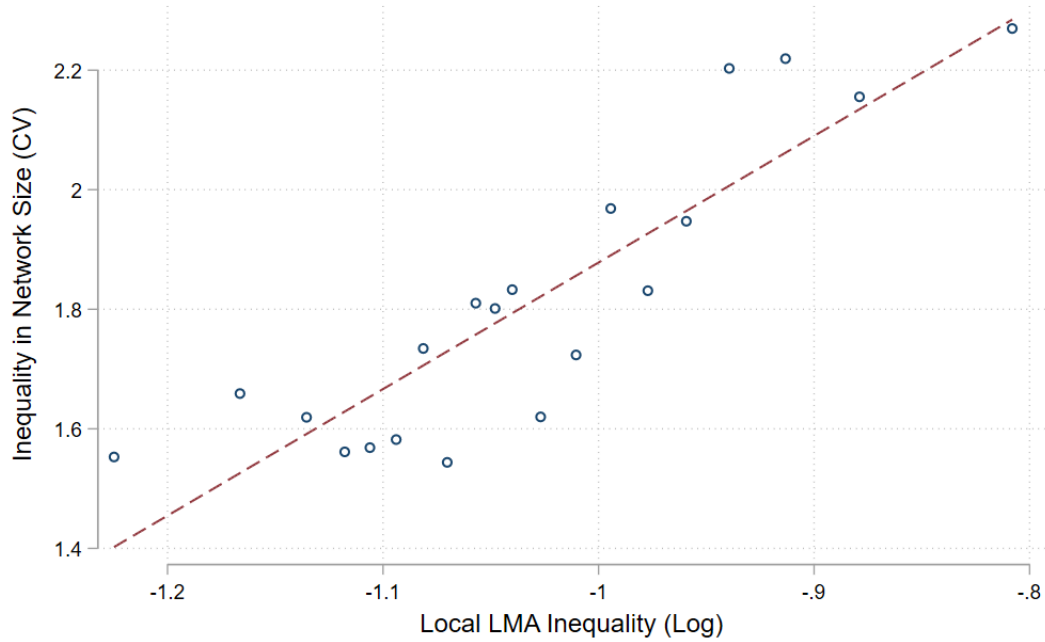
The slope is 0.04 with a standard error of 0.50.

Notes: This figure tests whether the Bartik-style instrument is relying on the variation that we want: occupations with growing national inequality driving increases in local inequality. In both panels, the vertical axis shows each occupation's impact on overall local inequality (estimates of $\hat{\pi}_{\kappa,t}$ from equation [20]), estimated at the occupation-by-year level. In Panel A, the horizontal axis shows growth in the occupation's inequality nationally, $\alpha_{\kappa,t}^{-1}$, relative to an omitted base occupation (electrical engineers). In Panel B, the horizontal axis shows growth in mean income within the upper tail of the income distribution for that occupation nationally. The strong positive relationship in Panel A indicates that the Bartik instrument indeed relies on occupations with growth in national inequality. Panel B shows that it is not relying on occupations that merely experience average income growth, as the relationship is flat.

Figure 9: Inequality and Physician Network Structure

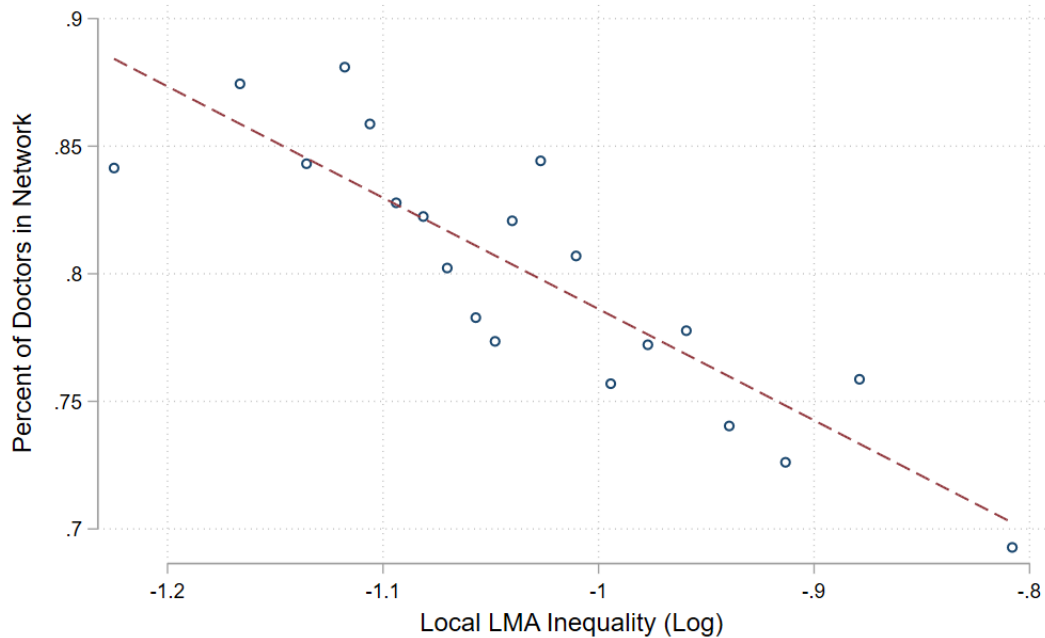
Panel A

Income inequality and inequality in MD network size



Panel B

Doctors' Extensive Margin Network Participation (ARF) and Income Inequality



Notes: Panel A shows the relationship between income inequality in an LMA (excluding physicians) and inequality in local ACA network size. Panel B shows the relationship between income inequality and the share of local physicians participating in any ACA network plan. In both cases, we group the data into twenty sized bins based on local income inequality. Sources: Income inequality is constructed from Census data provided by IPUMS (Ruggles et al. 2015). Physician network measures are based on authors' calculations from the University of Pennsylvania's Narrow Networks Project (Polsky and Weiner, 2015) and the Area Resource File.

Table 1: Descriptive Statistics for Top Occupations in 2000

	Mean Income	Inequality (α^{-1})	Occupation's share in:		
			top 10%	top 5%	top 1%
Chief executives and public administrators	\$265,885	0.66	0.05	0.08	0.14
Other financial specialists	178,688	0.60	0.01	0.02	0.03
Lawyers	246,985	0.59	0.03	0.04	0.06
Physicians	281,233	0.55	0.03	0.05	0.12
Financial managers	180,443	0.48	0.02	0.03	0.03
Supervisors and proprietors of sales jobs	127,309	0.44	0.04	0.04	0.04
Salespersons, n.e.c.	136,521	0.39	0.04	0.04	0.04
Accountants and auditors	121,396	0.38	0.02	0.02	0.02
Managers and administrators, n.e.c.	170,873	0.36	0.12	0.13	0.11
Managers and specialists in marketing, advertising, and public relations	170,156	0.36	0.04	0.04	0.04
Subject instructors (HS/college)	107,413	0.28	0.02	0.01	0.00
Production supervisors or foremen	95,536	0.26	0.01	0.01	0.00
Computer systems analysts and computer scientists	113,695	0.25	0.03	0.02	0.01
Not-elsewhere-classified engineers	128,842	0.24	0.01	0.01	0.00
Truck, delivery, and tractor drivers	70,855	0.24	0.01	0.01	0.01
Managers in education and related fields	112,626	0.24	0.01	0.01	0.00
Electrical engineer	133,983	0.23	0.01	0.01	0.00
Computer software developers	124,968	0.23	0.04	0.03	0.01
Registered nurses	79,703	0.21	0.02	0.01	0.00
Primary school teachers	66,338	0.17	0.01	0.00	0.00

Notes: This table shows basic descriptive statistics for the top twenty occupations in the top ten percent of the national income distribution in 2000. Column 1 reports mean income, where censored values have been replaced with the state-level mean income among those above the censoring point. The second column shows the occupation's income inequality, as measured with the inverse Pareto parameter. The final three columns show the occupation's share of all earners in the top ten, five, and one percent of the income distribution. Source: Authors' calculations using Census data from IPUMS (Ruggles et al., 2015).

Table 2: Wage income 1980-2014 for general population

Year	Median	percentile			p95/p90 (predicted)	p98/p95 (predicted)	$1/\alpha$
		p90	p95	p98			
1980	25.9	68.1	85.2	114.9	1.25 (1.26)	1.35 (1.35)	0.33
1990	29.0	76.1	96.8	137.6	1.27 (1.30)	1.42 (1.42)	0.38
2000	33.0	83.9	111.4	165.0	1.33 (1.32)	1.48 (1.44)	0.40
2014	30.5	90.0	120.0	177.9	1.33 (1.34)	1.48 (1.47)	0.42

Notes: Real wage income for observations with positive income (1000s of 2014 dollars using CPI). p95/p90 is the relative income of top 5 and top 10 per cent (predicted values in parentheses). Top-censoring prevents the calculation of 99th percentile wages. $1/\alpha$ are exponential parameters of a Pareto distribution and are calculated by MLE

Table 3: Wage income 1980-2014 for Physicians

Year	median	$1/\alpha$	p95/p90 (pred)
1980	100.6	1.29	2.44
1990	126.8	1.49	2.80
2000	137.5	1.58	2.99
2014	160.9	1.70	3.26

Notes: Real Wage Income (1000s of 2014 dollars using CPI). α is exponential parameter of a Pareto distribution and is calculated through MLE using the top 65 per cent of non-censored positive income

Table 4: Number of observations across labor market areas

Year	All				Physicians			
	Mean	25th	Median	75th	Mean	25th	Median	75th
1980	39584	17500	24560	37269	70	17	28	54
1990	43772	21207	29392	43286	89	22	37	70
2000	47541	21829	31786	46297	105	26	42	87
2014	51495	22789	34212	49489	138	29	56	115

Notes: Number of observations for labor market areas (all and physicians)

Table 5: Summary Table For Regression Variables

<i>Panel A: Logs</i>					
	Mean	SD	Min	Max	N
$\log(\alpha_o^{-1})$	0.28	0.50	-2.17	2.34	1012
$\log(\alpha_{-o}^{-1})$	-1.11	0.14	-1.58	-0.69	1012
Bartik Log	-1.14	0.10	-1.52	-0.90	1012
<i>Panel B: Levels</i>					
	Mean	SD	Min	Max	N
α_o^{-1}	1.48	0.76	0.11	10.38	1012
α_{-o}^{-1}	0.33	0.05	0.21	0.50	1012
Bartik Levels	0.33	0.02	0.26	0.39	1012

Notes: This table shows basic summary statistics for the variables in our Physician regressions. In both panels, $N = 1,012$. Source: authors' calculations using data from IPUMS (Ruggles et al., 2015).

Table 6: OLS Relationship for Physicians

	$\log(\alpha_o^{-1})$	$\log(\alpha_o^{-1})$	$\log(\alpha_o^{-1})$
$\log(\alpha_{-o}^{-1})$	0.240 (0.160)	0.216 (0.166)	0.249 (0.164)
Log Population		-0.029 (0.083)	-0.056 (0.081)
Log Average Income		0.126 (0.144)	0.373** (0.182)
N	1012	1012	1012
R^2	0.20	0.20	0.27
LMA FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Time-Varying Controls	No	No	Yes
Mean of Dep. Var.	0.28	0.28	0.28
SD of Dep. Var.	0.50	0.50	0.50
Mean of Indep. Var.	-1.11	-1.11	-1.11
SD of Indep. Var.	0.14	0.14	0.14

Notes: This table shows baseline OLS relationships between income inequality among physicians and other occupations in the top 20 among top income earners. Each column shows the raw OLS relationship between local income inequality among physicians and among the rest of the population in a given labor market area. Column 1 includes only area and year fixed effects. Column 2 adds controls for log population and log average wage income among those with positive income. Column 3 adds interactions between those controls and year fixed effects, allowing income and population to have differential effects by year. In all cases, inequality is measured as the log inverse of the Pareto parameter ($\log(\alpha^{-1})$). For physicians, we use incomes above the 35th percentile for this occupation to compute this Pareto parameter. The occupation of interest is denoted with o . Standard errors, in parenthesis, are clustered by labor market area. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Spillover Estimates for Physicians

	First Stage			Second Stage		
	$\log(\alpha_{-o}^{-1})$	$\log(\alpha_{-o}^{-1})$	$\log(\alpha_{-o}^{-1})$	$\log(\alpha_o^{-1})$	$\log(\alpha_o^{-1})$	$\log(\alpha_o^{-1})$
$\widehat{\log(\alpha_{-o}^{-1})}$				1.001** (0.410)	1.166** (0.516)	1.286** (0.524)
Bartik Log	1.152*** (0.344)	1.045*** (0.323)	0.949*** (0.312)			
Log Population		-0.049* (0.028)	-0.063** (0.027)		0.069 (0.107)	0.063 (0.101)
Log Average Income		0.067 (0.058)	0.045 (0.066)		0.074 (0.138)	0.334** (0.170)
N	1012	1012	1012	1012	1012	1012
R^2	0.82	0.83	0.83			
Kleibergen-Papp F				11.2	10.4	9.1
LMA FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Time-Varying Controls	No	No	Yes	No	No	Yes
Mean of Dep. Var.	-1.11	-1.11	-1.11	0.28	0.28	0.28
SD of Dep. Var.	0.14	0.14	0.14	0.50	0.50	0.50
Mean of Indep. Var.	-1.143	-1.143	-1.143	-1.107	-1.107	-1.107
SD of Indep. Var.	0.097	0.097	0.097	0.140	0.140	0.140

Notes: This table shows the full set of baseline results for income inequality spillovers to physicians from other occupations in the top 20 among top income earners. Columns 3, 4 and 5 show the first stage relationship between the Bartik instrument, using inequality among the top 20 non-physician occupations, and local non-physician income inequality. In these columns, the dependent variable is income inequality among non-physicians in an LMA. The Bartik instrument is defined in the text. Columns 4, 5 and 6 show the second stage of the two-stage least squares estimate using that same instrument. In these columns, the dependent variable is income inequality among physicians in an LMA, while non-physician inequality is the main right-hand-side variable. In columns 1 and 4, the only additional controls are location and year fixed effects. In columns 2 and 5, we add controls for log population and log average wage income among those with positive wage income as additional controls. Columns 3 and 6 add interactions between those controls and year fixed effects, allowing income and population to have differential effects by year. In all cases, inequality is measured as the log inverse of the Pareto parameter ($\log(\alpha^{-1})$). For physicians, we use incomes above the 35th percentile for this occupation to compute this Pareto parameter. The occupation of interest is denoted with o . Standard errors, in parenthesis, are clustered by labor market area. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Spillover Estimates for Dentists

	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	1st Stage	2SLS	2SLS
	$\log(1/\alpha(o))$	$\log(1/\alpha(o))$	$\log(1/\alpha(-o))$	$\log(1/\alpha(o))$	$\log(1/\alpha(o))$
$\log(1/\alpha(-o))$	0.60 [-0.12, 1.21]	0.54 [-0.16, 1.20]		2.17* [-0.11, 3.79]	2.77* [-0.19, 5.20]
Instrument			1.66*** [1.20, 2.04]		
Log of Population		-0.04 [-0.35, 0.35]	-0.12*** [-0.14,-0.08]		0.30 [-0.21, 0.83]
Log of Income		0.42 [-0.15, 0.85]	0.03 [-0.01, 0.08]		0.38 [-0.20, 0.87]
LMA FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
R^2 (ex. LMA FE)	0.19	0.20	0.87	.	.
Observations	160	160	160	160	160

Notes: This table shows the full set of baseline results for income inequality spillovers to dentists from other occupations in the top 20 among top income earners. Columns 1 and 2 show the OLS relationship between local income inequality among dentists and among the rest of the population in a given labor market area. Column 3 shows the first stage relationship between the Bartik instrument, using inequality among the top 20 non-dentist occupations, and local non-dentist income inequality. Columns 4 and 5 show the second stage of the two-stage least squares estimate using that same instrument. In column 3 only, the dependent variable is income inequality among non-dentists in an LMA. In the remaining columns, the dependent variable is income inequality among dentists in an LMA, while non-dentist inequality is the main right-hand-side variable. In all cases, inequality is measured as the log inverse of the Pareto parameter ($\log(1/\alpha)$). For dentists, we use incomes above the 35th percentile for this occupation to compute this Pareto parameter. The occupation of interest is denoted with o . Standard errors are bootstrapped using 300 draws, stratified at the occupation-year-labor market level. Square brackets show the 95 percent confidence interval based on this bootstrapping. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The income control is log average wage income for those with positive income.

Table 9: Spillover Estimates for Real Estate Agents

	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	1st Stage	2SLS	2SLS
	$\log(1/\alpha(o))$	$\log(1/\alpha(o))$	$\log(1/\alpha(-o))$	$\log(1/\alpha(o))$	$\log(1/\alpha(o))$
$\log(1/\alpha(-o))$	0.17*	0.17*		1.02**	1.32**
	[-0.03, 0.30]	[-0.03, 0.30]		[0.20, 2.09]	[0.29, 2.56]
Instrument			0.64***		
			[0.51, 0.76]		
Log of Population		0.05*	-0.04***		0.11***
		[-0.00, 0.10]	[-0.05,-0.03]		[0.04, 0.20]
Log of Income		0.23***	0.03**		0.19**
		[0.13, 0.33]	[0.00, 0.05]		[0.08, 0.31]
LMA FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
R^2 (ex. LMA FE)	0.05	0.05	0.82	.	.
Observations	1,448	1,448	1,448	1,448	1,448

Notes: This table shows the full set of baseline results for income inequality spillovers to real estate agents from other occupations in the top 20 among top income earners. Columns 1 and 2 show the OLS relationship between local income inequality among realtors and among the rest of the population in a given labor market area. Column 3 shows the first stage relationship between the Bartik instrument, using inequality among the top 20 non-realtor occupations, and local non-realtor income inequality. Columns 4 and 5 show the second stage of the two-stage least squares estimate using that same instrument. In column 3 only, the dependent variable is income inequality among non-realtors in an LMA. In the remaining columns, the dependent variable is income inequality among realtors in an LMA, while non-realtor inequality is the main right-hand-side variable. In all cases, inequality is measured as the log inverse of the Pareto parameter ($\log(1/\alpha)$). For real estate agents, we use incomes above the 80th percentile for this occupation to compute this Pareto parameter. The occupation of interest is denoted with o . Standard errors are bootstrapped using 300 draws, stratified at the occupation-year-labor market level. Square brackets show the 95 percent confidence interval based on this bootstrapping. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The income control is log average wage income for those with positive income.

Table 10: Spillover Estimates for Financial Managers

	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	1st Stage	2SLS	2SLS
	$\log(1/\alpha(o))$	$\log(1/\alpha(o))$	$\log(1/\alpha(-o))$	$\log(1/\alpha(o))$	$\log(1/\alpha(o))$
$\log(1/\alpha(-o))$	0.77*** [0.21, 1.34]	0.63** [0.01, 1.26]		1.29** [0.15, 2.64]	1.02 [-0.81, 2.78]
$\log(I)$			1.27*** [1.20, 1.30]		
\logpop		-0.21 [-0.50, 0.13]	-0.08*** [-0.09,-0.07]		-0.16 [-0.61, 0.23]
$\log(\text{income})$		0.13 [-0.27, 0.56]	0.01 [-0.00, 0.01]		0.14 [-0.26, 0.58]
LMA FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
R^2 (ex. LMA FE)	0.46	0.47	0.89	.	.
Observations	184	184	184	184	184

Notes: This table shows the full set of baseline results for income inequality spillovers to financial managers from other occupations in the top 20 among top income earners. Columns 1 and 2 show the OLS relationship between local income inequality among financial managers and among the rest of the population in a given labor market area. Column 3 shows the first stage relationship between the Bartik instrument, using inequality among the top 20 non-financier occupations, and local non-financier income inequality. Columns 4 and 5 show the second stage of the two-stage least squares estimate using that same instrument. In column 3 only, the dependent variable is income inequality among non-financiers in an LMA. In the remaining columns, the dependent variable is income inequality among financial managers in an LMA, while non-financier inequality is the main right-hand-side variable. In all cases, inequality is measured as the log inverse of the Pareto parameter ($\log(1/\alpha)$). For financial managers, we use incomes above the 90th percentile for this occupation to compute this Pareto parameter. The occupation of interest is denoted with o . Standard errors are bootstrapped using 300 draws, stratified at the occupation-year-labor market level. Square brackets show the 95 percent confidence interval based on this bootstrapping. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The income control is log average wage income for those with positive income.

Table 11: Spillover Estimates for Nurses

	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	1st Stage	2SLS	2SLS
	$\log(1/\alpha(o))$	$\log(1/\alpha(o))$	$\log(1/\alpha(-o))$	$\log(1/\alpha(o))$	$\log(1/\alpha(o))$
$\log(1/\alpha(-o))$	0.37 [-0.14, 0.75]	0.32 [-0.17, 0.77]		0.07 [-1.09, 1.14]	-0.11 [-1.64, 1.46]
$\log(I)$			1.09*** [1.08, 1.12]		
\logpop		-0.07 [-0.27, 0.11]	-0.10*** [-0.10,-0.10]		-0.14 [-0.46, 0.16]
$\log(\text{income})$		0.36*** [0.09, 0.65]	-0.00 [-0.00, 0.00]		0.35** [0.07, 0.64]
LMA FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
R^2 (ex. LMA FE)	0.59	0.61	0.89	.	.
Observations	184	184	184	184	184

Notes: This table shows the full set of baseline results for income inequality spillovers to nurses from other occupations in the top 20 among top income earners. Columns 1 and 2 show the OLS relationship between local income inequality among nurses and among the rest of the population in a given labor market area. Column 3 shows the first stage relationship between the Bartik instrument, using inequality among the top 20 non-nursing occupations, and local non-nursing income inequality. Columns 4 and 5 show the second stage of the two-stage least squares estimate using that same instrument. In column 3 only, the dependent variable is income inequality among non-realtors in an LMA. In the remaining columns, the dependent variable is income inequality among nurses in an LMA, while non-nurse inequality is the main right-hand-side variable. In all cases, inequality is measured as the log inverse of the Pareto parameter ($\log(1/\alpha)$). For nurses, we use incomes above the 90th percentile for this occupation to compute this Pareto parameter. The occupation of interest is denoted with o . Standard errors are bootstrapped using 300 draws, stratified at the occupation-year-labor market level. Square brackets show the 95 percent confidence interval based on this bootstrapping. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The income control is log average wage income for those with positive income.

Table 12: Spillover Estimates for College Professors

	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	1st Stage	2SLS	2SLS
	$\log(1/\alpha(o))$	$\log(1/\alpha(o))$	$\log(1/\alpha(-o))$	$\log(1/\alpha(o))$	$\log(1/\alpha(o))$
$\log(1/\alpha(-o))$	0.18 [-0.31, 0.65]	0.34 [-0.13, 0.89]		0.44 [-0.82, 1.53]	0.73 [-0.77, 2.23]
$\log(I)$			1.28*** [1.27, 1.31]		
logpop		0.22*** [0.04, 0.46]	-0.07*** [-0.07,-0.07]		0.27** [0.01, 0.59]
$\log(\text{income})$		-0.21 [-0.50, 0.16]	0.07*** [0.07, 0.08]		-0.22 [-0.52, 0.15]
LMA FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
R^2 (ex. LMA FE)	0.70	0.71	0.89	.	.
Observations	184	184	184	184	184

Notes: This table shows the full set of baseline results for income inequality spillovers to professors from other occupations in the top 20 among top income earners. Columns 1 and 2 show the OLS relationship between local income inequality among professors and among the rest of the population in a given labor market area. Column 3 shows the first stage relationship between the Bartik instrument, using inequality among the top 20 non-professor occupations, and local non-professor income inequality. Columns 4 and 5 show the second stage of the two-stage least squares estimate using that same instrument. In column 3 only, the dependent variable is income inequality among non-professors in an LMA. In the remaining columns, the dependent variable is income inequality among professors in an LMA, while non-professor inequality is the main right-hand-side variable. In all cases, inequality is measured as the log inverse of the Pareto parameter ($\log(1/\alpha)$). For professors, we use incomes above the 90th percentile for this occupation to compute this Pareto parameter. The occupation of interest is denoted with o . Standard errors are bootstrapped using 300 draws, stratified at the occupation-year-labor market level. Square brackets show the 95 percent confidence interval based on this bootstrapping. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The income control is log average wage income for those with positive income.

Table 13: Spillover Estimates for Secretaries

	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	1st Stage	2SLS	2SLS
	$\log(1/\alpha(o))$	$\log(1/\alpha(o))$	$\log(1/\alpha(-o))$	$\log(1/\alpha(o))$	$\log(1/\alpha(o))$
$\log(1/\alpha(-o))$	0.03 [-0.11, 0.21]	0.06 [-0.09, 0.24]		0.39 [-0.69, 1.38]	0.41 [-0.68, 1.30]
Instrument			0.69*** [0.57, 0.81]		
Log of Population		0.13*** [0.05, 0.22]	-0.06*** [-0.07,-0.04]		0.16*** [0.04, 0.27]
Log of Income		0.16*** [0.03, 0.31]	0.02*** [0.00, 0.04]		0.15** [0.01, 0.31]
LMA FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
R^2 (ex. LMA FE)	0.13	0.14	0.80	.	.
Observations	1,576	1,576	1,576	1,576	1,576

Notes: This table shows the full set of baseline results for income inequality spillovers to secretaries from other occupations in the top 20 among top income earners. Columns 1 and 2 show the OLS relationship between local income inequality among secretaries and among the rest of the population in a given labor market area. Column 3 shows the first stage relationship between the Bartik instrument, using inequality among the top 20 non-secretary occupations, and local non-secretary income inequality. Columns 4 and 5 show the second stage of the two-stage least squares estimate using that same instrument. In column 3 only, the dependent variable is income inequality among non-secretaries in an LMA. In the remaining columns, the dependent variable is income inequality among secretaries in an LMA, while non-secretary inequality is the main right-hand-side variable. In all cases, inequality is measured as the log inverse of the Pareto parameter ($\log(1/\alpha)$). For secretaries, we use incomes above the 90th percentile for this occupation to compute this Pareto parameter. The occupation of interest is denoted with o . Standard errors are bootstrapped using 300 draws, stratified at the occupation-year-labor market level. Square brackets show the 95 percent confidence interval based on this bootstrapping. Statistical significance is denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The income control is log average wage income for those with positive income.

Table 14: Inequality and Physician Network Structure

	(1)	(2)	(3)	(4)
Dependent variable (z -scores):	SD network size	SD network size	Network participation	Network participation
$\log(\alpha^{-1})$ for non-physicians (z -score)	0.724*** (0.110)	0.616*** (0.114)	-0.420*** (0.070)	-0.225*** (0.068)
N	253	253	253	253
R^2	0.25	0.31	0.10	0.36
SD of Dep. Var.	258.47	258.47	0.56	0.56
SD of $\log(\alpha^{-1})$	0.14	0.14	0.14	0.14
Controls (specialty comp.)	No	Yes	No	Yes

Notes: This table shows cross-sectional regressions of network inequality measures against local income inequality. The dependent variable in columns 1 and 2 is the standard deviation of the number of physicians in local ACA networks. In columns 3 and 4, it is the share of local physicians participating in any local ACA exchange plan network. We standardize both the left- and right-hand-side variables for ease of interpretation. The standard deviations of the original (non-standardized) variables are provided in the table. Sources: Income inequality is constructed from Census data provided by IPUMS (Ruggles et al. 2015). Physician network data are based on authors' calculations from the University of Pennsylvania's Narrow Networks Project (Polsky and Weiner, 2015) and the Area Resource File.

Table 15: Inequality and Physician Pricing Dispersion

Dependent variable (prices):	(1)	(2)	(3)	(4)	(5)	(6)
	$\Delta \log \alpha^{-1}$	$\Delta \log \alpha^{-1}$	$\Delta 75/50$ ratio	$\Delta 90/50$ ratio	$\log \alpha_{prices}^{-1}$	$\log \alpha_{prices}^{-1}$
$\Delta \log(I)$	1.641*	3.423*	0.962**	2.723		
	(0.797)	(1.619)	(0.441)	(2.378)		
$\log(\alpha_{incomes}^{-1})$					0.746***	0.516**
					(0.251)	(0.243)
Log population					-0.099***	-0.184***
					(0.016)	(0.029)
Log income						1.190***
						(0.342)
N	20	20	20	20	79	79
R^2	0.06	0.34	0.16	0.02	0.89	0.91
SD of Dep. Var.	0.112	0.112	0.039	0.283	0.517	0.517
SD of Indep. Var.	0.016	0.016	0.016	0.016	0.115	0.115
States	TX	TX	TX	TX	TX,CO,NH	TX,CO,NH
Controls (specialty comp.)	No	Yes	No	No	No	No

Notes: This table shows short-difference and cross-sectional regressions of inequality in physician reimbursements against local income inequality. Columns 1-4 use BCBS-TX pricing data and short differences, while columns 5 and 6 add in Colorado and New Hampshire APCD data and treat the data as a cross-section. The dependent variable in columns 1, 2, 5, and 6 is our standard inequality measure: $\log(\alpha^{-1})$ but calculated using physician markups. Columns 3 and 4 use the 75/50 and 90/50 ratio of physician markups, respectively. The independent variable in columns 1-4 is the change in the value of the Bartik instrument from 2000-2014 (still using as weights the local occupational distribution in 1980), while in columns 5 and 6 we use the 2014 local inequality (excluding physicians). The geographic unit in columns 1-4 is Labor Market Areas in Texas, and in columns 5-6 is Commuting Zones in all three states. Sources: Income inequality is constructed from Census data provided by IPUMS (Ruggles et al. 2015). Physician pricing inequality measures are based on authors' calculations from BCBS-TX, APCD-CO, and APCD-NH data.

A Proofs of main results

A.1 Positive assortative matching in equilibrium

Here we show that the equilibrium must feature positive assortative matching between the income of the patient and the skill of the doctor. To do so, we assume that there are 2 individuals 1 and 2 with income $x_1 < x_2$ whose consumption bundles are so that $z_1 > z_2$ and $c_1 < c_2$. For simplicity we write the utility function as a function of health services and the income left for other goods ($x - \omega(z)$).

Note that since consumer 1 chooses a doctor of quality z_1 , it must be the case that:

$$u(z_1, x_1 - \omega(z_1)) \geq u(z_2, x_1 - \omega(z_2)).$$

Further, we have:

$$\begin{aligned} & u(z_1, x_2 - \omega(z_1)) - u(z_2, x_2 - \omega(z_2)) \\ &= u(z_1, x_2 - \omega(z_1)) - u(z_1, x_1 - \omega(z_1)) + u(z_1, x_1 - \omega(z_1)) - u(z_2, x_1 - \omega(z_2)) \\ &\quad + u(z_2, x_1 - \omega(z_2)) - u(z_2, x_2 - \omega(z_2)) \\ &= \int_{x_1 - \omega(z_1)}^{x_2 - \omega(z_1)} \left(\frac{\partial u}{\partial c}(z_1, c) - \frac{\partial u}{\partial c}(z_2, c) \right) + u(z_1, x_1 - \omega(z_1)) - u(z_2, x_1 - \omega(z_2)). \end{aligned}$$

If the utility function has a positive cross-partial (which is the case for a Cobb-Douglas), then the first term is positive as $z_1 > z_2$. Since the second term is also weakly positive, then it must be the case that $u(z_1, x_2 - \omega(z_1)) > u(z_2, x_2 - \omega(z_2))$, in other words, consumer 2 would rather pick a doctor of ability z_1 . Therefore there is a contradiction and it must be the case that $z_1 < z_2$.

A.2 Solving (6)

We look for a specific solution to(6) of the type $w(z) = K_1 z^{\frac{\alpha_z}{\alpha_x}}$. We find that such a K_1 must satisfy

$$K_1 = x_{\min} \frac{\beta_z \alpha_x \lambda}{\alpha_z (1 - \beta_z) + \beta_z \alpha_x} \left(\frac{1}{z_c} \right)^{\frac{\alpha_z}{\alpha_x}}.$$

As the solutions to the differential equation $w'(z)z + \frac{\beta_z}{1-\beta_z}w(z) = 0$ are given by $Kz^{-\frac{\beta_z}{1-\beta_z}}$ for any constant K . We get that all solutions to (6) take the form:

$$w(z) = \frac{x_{\min}\beta_z\alpha_x\lambda}{\alpha_z(1-\beta_z) + \beta_z\alpha_x} \left(\frac{z}{z_c}\right)^{\frac{\alpha_x}{\alpha_z}} + Kz^{-\frac{\beta_z}{1-\beta_z}}.$$

We then obtain (7) by using that $w(z_c) = x_{\min}$ which fixes

$$K = x_{\min}z_c^{\frac{\beta_z}{1-\beta_z}} \frac{\alpha_z(1-\beta_z) + \beta_z\alpha_x(1-\lambda)}{\alpha_z(1-\beta_z) + \beta_z\alpha_x}.$$

A.3 Proof of Proposition 2

Using (1), (2), (5) and (10), we get that the utility of a consumer with income x is given by

$$\begin{aligned} u(x) &= (x - h(x))^{1-\beta_z} (m^{-1}(x))^{\beta_z} \\ &= \left(\frac{\alpha_z(1-\beta_z)}{\alpha_z(1-\beta_z) + \beta_z\alpha_x} x - \frac{1}{\lambda} \frac{\alpha_z(1-\beta_z) + \beta_z\alpha_x(1-\lambda)}{\alpha_z(1-\beta_z) + \beta_z\alpha_x} x_{\min} \left(\frac{x}{x_{\min}}\right)^{-\frac{\alpha_x}{\alpha_z} \frac{\beta_z}{1-\beta_z}} \right) \\ &\quad \cdot \left(z_c \left(\frac{x}{x_{\min}}\right)^{\frac{\alpha_x}{\alpha_z}} \right)^{\beta_z}. \end{aligned}$$

Therefore $eq(x)$ obeys

$$\begin{aligned} eq(x) &= \left(\frac{\alpha_z(1-\beta_z)}{\alpha_z(1-\beta_z) + \beta_z\alpha_x} x - \frac{1}{\lambda} \frac{\alpha_z(1-\beta_z) + \beta_z\alpha_x(1-\lambda)}{\alpha_z(1-\beta_z) + \beta_z\alpha_x} x_{\min} \left(\frac{x}{x_{\min}}\right)^{-\frac{\alpha_x}{\alpha_z} \frac{\beta_z}{1-\beta_z}} \right) \\ &\quad \cdot \left(\frac{x}{x_{\min}}\right)^{\frac{\alpha_x}{\alpha_z} \frac{\beta_z}{1-\beta_z}}, \end{aligned}$$

which implies that for x large enough

$$eq(x) \approx \frac{\alpha_z(1-\beta_z)x_{\min}^{-\frac{\alpha_x}{\alpha_z} \frac{\beta_z}{1-\beta_z}}}{\alpha_z(1-\beta_z) + \beta_z\alpha_x} x^{1 + \frac{\alpha_x}{\alpha_z} \frac{\beta_z}{1-\beta_z}}.$$

Then the distribution of real income obeys $\Pr(EQ > e) = \Pr(X > eq^{-1}(e))$, so that for e large enough, we obtain:

$$\Pr(EQ > e) \approx \left(\frac{x_{\min}\alpha_z(1-\beta_z)}{\alpha_z(1-\beta_z) + \beta_z\alpha_x} \frac{1}{e} \right)^{\frac{\alpha_x}{\alpha_z} \frac{\beta_z}{1-\beta_z}}.$$

Therefore asymptotically, real income is distributed in a Pareto way with a shape parameter $\alpha_{eq} \equiv \frac{\alpha_x}{1 + \frac{\alpha_x}{\alpha_z} \frac{\beta_z}{1 - \beta_z}}$. Moreover we obtain: $\frac{d \ln \alpha_{eq}}{d \ln \alpha_x} = \frac{1}{1 + \frac{\alpha_x}{\alpha_z} \frac{\beta_z}{1 - \beta_z}}$.

B Formalizing the extensions

B.1 Adding Brewers: the Role of Assortative Matching

Taking first order conditions with respect to c and y , we obtain that expenditures on beers and on the homogeneous good are related by

$$py = \frac{\beta_y}{1 - \beta_y - \beta_z} c. \quad (22)$$

The first order condition with respect to the quality of the health services consumed and the homogeneous good similarly imply

$$\omega'(z) z = \frac{\beta_z}{1 - \beta_y - \beta_z} c. \quad (23)$$

Together with the budget constraint equation

$$\omega(z) + py + c = x,$$

(22) and (23) give (3) so that all results concerning $w(z)$ including (10) still apply, and

$$y(x) = \frac{1}{p} \frac{\beta_y}{1 - \beta_z} (x - h(x)). \quad (24)$$

Market clearing imposes

$$\int_{x_{\min}}^{\infty} y(x) dG_x(x) = \mu_m \int_{y_c}^{\infty} y dG_y(y), \quad (25)$$

where $y(x)$ denotes the consumption of beer by a consumer of income x and G_a the cdf of

variable a . Plugging (24) in (25) we obtain:

$$py_c = \frac{\psi}{\mu_m} \left(\frac{y_c}{y_{\min}} \right)^{\alpha_y},$$

$$\text{with } \psi \equiv \frac{\alpha_y - 1}{\alpha_y} \frac{\alpha_z \beta_y \left(\frac{1}{\alpha_x} + \lambda - 1 \right)}{\lambda (\beta_z + \alpha_z (1 - \beta_z))} \widehat{x}.$$

This implies that there are two possible scenarios. If $\psi \geq \mu_m x_{\min}$ then $py_{\min} \geq x_{\min}$ so that all possible brewers end up working as brewers. We then have

$$p = \frac{\psi}{\mu_m y_{\min}}.$$

Since ψ is decreasing in α_x , a decrease in the shape parameter of consumer income is associated with a proportional increase in brewer's income.

Note that

$$\frac{\psi}{x_{\min}} = \frac{\alpha_z \beta_y (\alpha_y - 1)}{\lambda (\beta_z + \alpha_z (1 - \beta_z)) \alpha_y} \frac{1 + (\lambda - 1) \alpha_x}{\alpha_x - 1}$$

is decreasing in α_x . Therefore as α_x decreases then this situation becomes more and more likely.

Otherwise, $y_c > y_{\min}$ with

$$y_c = y_{\min} \left(\mu_m \frac{x_{\min}}{\psi} \right)^{\frac{1}{\alpha_y}},$$

so that as α_x decreases (and consequently x_{\min} to keep mean income of consumers constant), y_c decreases and more and more potential brewers decide to become brewers. This leads to

$$p = \left(\frac{\psi}{\mu_m} \right)^{\frac{1}{\alpha_y}} \left(\frac{\alpha_x - 1}{\alpha_x} \widehat{x} \right)^{\frac{\alpha_y - 1}{\alpha_y}} \frac{1}{y_{\min}}$$

$$= \left(\frac{1}{\mu_m} \frac{\alpha_z \beta_y (\alpha_y - 1)}{\lambda (\beta_z + \alpha_z (1 - \beta_z)) \alpha_y} \right)^{\frac{1}{\alpha_y}} \left(\frac{1 + (\lambda - 1) \alpha_x}{\alpha_x - 1} \right)^{\frac{1}{\alpha_y}} \frac{\alpha_x - 1}{\alpha_x} \frac{\widehat{x}}{y_{\min}}.$$

Note that

$$\frac{d}{d\alpha_x} (1 + (\lambda - 1) \alpha_x)^{\frac{1}{\alpha_y}} \frac{(\alpha_x - 1)^{\frac{\alpha_y - 1}{\alpha_y}}}{\alpha_x}$$

$$= [(1 + (\lambda - 1) \alpha_x) (\alpha_y - 1) - (\alpha_x - 1)] \frac{(1 + (\lambda - 1) \alpha_x)^{\frac{1}{\alpha_y} - 1} (\alpha_x - 1)^{\frac{-1}{\alpha_y}}}{\alpha_y \alpha_x^2},$$

the sign of which is ambiguous since λ can be close to 1 and we may have $\alpha_x > \alpha_y$. Therefore in this case, a decrease in α_x increases the supply of beers but as a result the impact on brewers' income is ambiguous.

For any price level \tilde{p} , we can define the real welfare measure similarly as the income which gives the same utility in the market and when the agent is forced to consume (for free) z_c while having y prices at \tilde{p} . That is we now have:

$$u\left(z_c, \frac{1 - \beta_z - \beta_y}{1 - \beta_z} eq(x), \frac{\beta_y}{1 - \beta_z} \frac{eq(x)}{\tilde{p}}\right) = u(z(x), c(x), y(x)).$$

We then obtain:

$$eq(x) = \left(\frac{\tilde{p}}{p}\right)^{\frac{\beta_y}{1-\beta_z}} \left(\frac{\alpha_z(1-\beta_z)}{\alpha_z(1-\beta_z) + \beta_z\alpha_x} x - \frac{1}{\lambda} \frac{\alpha_z(1-\beta_z) + \beta_z\alpha_x(1-\lambda)}{\alpha_z(1-\beta_z) + \beta_z\alpha_x} x_{\min} \left(\frac{x}{x_{\min}}\right)^{-\frac{\alpha_x}{\alpha_z} \frac{\beta_z}{1-\beta_z}} \right) \left(\frac{x}{x_{\min}}\right)^{\frac{\alpha_x}{\alpha_z} \frac{\beta_z}{1-\beta_z}}.$$

Therefore the analysis of real income inequality is the same whether $\beta_y = 0$ or not.

B.2 Occupational Mobility

Above we assumed that a potential doctor working as a consumer makes the minimum amount possible as a consumer: x_{min} . In reality it is quite plausible that those succeeding as doctors would have succeeded in other occupations as well (Kirkeboen, Leuven and Mogstad, 2016). To capture this, we now switch to the opposite extreme and assume that there is perfect correlation between abilities as a doctor and as a consumer. We keep the model as before, except we assume that there is a mass 1 of agents who decide whether they want to be doctors or consumers. We rank agents in descending order of ability and use i to denote their rank, so that the most skilled agent has rank 0 and the least skilled has rank 1. For two agents i and i' with $i < i'$, i will be better both as a consumer and as a doctor than i' . We assume that both ability distributions are Pareto with parameters (x_{min}, α_x) for consumer and (z_{min}, α_z) for doctors. An agent i can choose between becoming a consumer earning $x(i)$ or being a doctor providing health services of quality $z(i)$ and earning $w(z(i))$. Those working as doctors also need the services of doctors. We assume that $\lambda > 1$ to ensure that everyone can get health services. By definition of the rank we have that the

counter-cumulative distribution functions (1 minus the CDFs) for x and z obey:

$$\overline{G}_x(x(i)) = \overline{G}_z(z(i)) = i.$$

In equilibrium, it is always the case that below a certain rank, some individuals will choose to be doctors. In addition under parameter conditions detailed below, some individuals will also choose to be consumers (details in Appendix B.2.1). That is, for i low enough, agents must be indifferent between becoming a doctor or a consumer: $w(z(i)) = x(i)$, which directly implies that, for z high enough, the wage function must satisfy:

$$w(z) = \overline{G}_x^{-1}(\overline{G}_z(z)).$$

Since both ability distributions are Pareto, this can be written as:

$$w(z) = x_{\min} \left(\frac{z}{z_{\min}} \right)^{\frac{\alpha_z}{\alpha_x}}. \quad (26)$$

Doctor wages grow in proportion to what they could earn as a consumer.

Let $\mu(z) \in (0, 1)$ denotes the share of individuals able to provide health services of quality z who are doctors. For z sufficiently high that individuals of rank $\overline{G}_z(z)$ and below and their patient work both as consumer and doctors, market clearing implies

$$\left(\frac{x_{\min}}{m(z)} \right)^{\alpha_x} = \int_z^\infty \lambda \mu(\zeta) g_z(\zeta) d\zeta, \quad (27)$$

where $m(z)$ denotes the income (earned either as a consumer or a doctor) of the patient of a doctor of quality z .

The first order condition on health care consumption (3) still applies, and together with (26) and (27) it implies that:

$$\int_z^\infty \mu(\zeta) \alpha_z \zeta^{-\alpha_z - 1} d\zeta = \lambda^{\alpha_x - 1} z^{-\alpha_z} \left(\frac{\alpha_z}{\alpha_x} + \frac{\beta_z}{1 - \beta_z} \right)^{-\alpha_x}.$$

Differentiating with respect to z , we find that μ is a constant: $\mu = \lambda^{\alpha_x - 1} \left(\frac{\alpha_z}{\alpha_x} + \frac{\beta_z}{1 - \beta_z} \right)^{-\alpha_x}$. Intuitively, with a constant μ , doctors' wages grow proportionately with patients' incomes, in line with the Cobb-Douglas assumption. Note that since we assumed that $\mu < 1$, this situation is only possible as long as $\lambda^{\alpha_x - 1} \left(\frac{\alpha_z}{\alpha_x} + \frac{\beta_z}{1 - \beta_z} \right)^{-\alpha_x} < 1$.

Therefore, if $\lambda^{\alpha_x - 1} \left(\frac{\alpha_z}{\alpha_x} \frac{1 - \beta_z}{\beta_z} + 1 \right)^{-\alpha_x} < 1$, we have that $P_{doc}(W_d > w_d) = P(Z > w^{-1}(w_d))$ for w_d high enough so that the observed distribution for doctor wages is Pareto with a shape parameter α_x : Proposition 1 still applies (in fact the distribution is now exactly Pareto above a threshold). We solve for the full model in Appendix B.2.1.²⁹ Further, if the distributions of x and z are only asymptotically Pareto, then our results remain true asymptotically, so that Proposition 1 applies.

Note that in terms of observed top income inequality the model where agents can switch and the one where they cannot are observationally equivalent: doctors' top income inequality perfectly traces that of the consumers. This is so because even when doctors are not allowed to shift across occupations, the relative reward to the very best doctors adjusts correspondingly with the shift for consumers.

Supply versus demand side effects. In the model just presented doctors and consumers interact both through a demand effect—consumers are the clients of doctors—and a supply effect—doctors can choose to become consumers. Since the wage level is directly determined by doctors' outside option (according to (26)), one may think that the mechanism which leads to spillovers in income inequality is very different compared to the demand-side mechanism of the baseline model. This is, however, not the case. In Appendix B.2.2 we split the role of consumers into two: patients, who only serve the role of consumers of doctor services and an “outside option” which only serves the role of providing doctors with an alternative occupation to providing medical services. We show that when the utility function is given by (1), the income inequality of doctors is entirely driven by that of their patients and is independent of changes in the income inequality for the outside option. Consequently, the driving force is still the demand side.

B.2.1 Different adjustment margin

In this appendix we fully solve the model described in section 2.2.2. First note that above a certain threshold, there will be individuals choosing to be doctors. Assume that this is not the case, then there is an upper bound z_M on the quality of health care provided. Consider an individual 1 with income X who is a consumer. Her utility obeys $u(X) \leq z_M^{\beta_z} X^{1 - \beta_z}$. Consider now individual 2 with consumer ability $X^{\frac{1}{2}}$. For X large enough, this individual would be a consumer. Assume, however, that she switches and decides to become a doctor,

²⁹If $\lambda^{\alpha_x - 1} \left(\frac{\alpha_z}{\alpha_x} \frac{1 - \beta_z}{\beta_z} + 1 \right)^{-\alpha_x} > 1$, then all individuals above a certain ability threshold choose to be doctors while all those below it choose to be consumers. This seems counterfactual.

then she would provide health care service with quality $z_{\min} \left(\frac{X^{\frac{1}{2}}}{x_{\min}} \right)^{\frac{\alpha_x}{\alpha_z}}$. Individual 1 would then rather hire individual 2 as a doctor and consume $\frac{1}{2}X$ in homogeneous good. Under this alternative allocation her utility is $\left(\frac{1}{2}\right)^{1-\beta_z} X^{1-\beta_z} z_{\min}^{\beta_z} \left(\frac{X^{\frac{1}{2}}}{x_{\min}}\right)^{\frac{\alpha_x}{\alpha_z} \beta_z}$, which for X high enough is higher than the utility under the original allocation. Individual 2 earns $\frac{1}{2}X$ which is also higher than her initial income. Therefore this is a profitable deviation and the initial allocation cannot be an equilibrium.

As a result, the equilibrium must be that below a certain rank some individuals choose to be doctors. We then have 3 possible cases, which we will solve in turn:

- Below a certain rank individuals choose to be both doctors and consumers and above it they all choose to be consumers;
- Below a certain rank individuals choose to be both doctors and consumers and above it they all choose to be doctors;
- Below a certain rank, all individuals choose to be doctors.

Case 1. Consider first the case where there exists a z_c such that individuals of rank higher than $\bar{G}_z(z_c)$ all choose to be consumers. Then (27) applies for $z > z_c$ and we know that for $z \geq z_c$, $\mu = \frac{\lambda^{\alpha_x - 1}}{\left(\frac{\alpha_z}{\alpha_x} \frac{1 - \beta_z}{\beta_z} + 1\right)^{\alpha_x}}$, which we assume to be smaller than 1. Since $m(z_c) = x_{\min}$, we obtain:

$$z_c = z_{\min} \left(\frac{\lambda}{\frac{\alpha_z}{\alpha_x} \frac{1 - \beta_z}{\beta_z} + 1} \right)^{\frac{\alpha_x}{\alpha_z}}, \quad (28)$$

which is only possible if $\lambda \geq \frac{\alpha_z}{\alpha_x} \frac{1 - \beta_z}{\beta_z} + 1$.

Case 2. Consider now the opposite case. Individuals ranked above $\bar{G}_z(z_m)$ all choose to be consumers, those ranked below are indifferent. Since $\lambda > 1$, the supply of health services by agents ranked higher than $\bar{G}_z(z_m)$ is enough to cover their own demand for health services. Therefore, if one denotes by $r(z)$ the rank of the patient of a doctor of quality z , we obtain that there exists a $z_p < z_m$, such that $r(z_p) = z_m$: doctors with ability lower than z_p only provide health services to doctors and those with ability above z_p provide health services to both doctors and consumers. Since $z_m > z_p$, we have that for $z \geq z_m$, (27) applies which directly leads to $\mu = \frac{\lambda^{\alpha_x - 1}}{\left(\frac{\alpha_z}{\alpha_x} \frac{1 - \beta_z}{\beta_z} + 1\right)^{\alpha_x}}$ for $z \geq z_m$. This imposes, as before, the restriction $\frac{\lambda^{\alpha_x - 1}}{\left(\frac{\alpha_z}{\alpha_x} \frac{1 - \beta_z}{\beta_z} + 1\right)^{\alpha_x}} < 1$. We then get to further write for $z \leq z_m$:

$$r(z) = \int_z^{z_m} \lambda g_z(\zeta) d\zeta + \int_{z_m}^{\infty} \lambda \mu(\zeta) g_z(\zeta) d\zeta = \lambda \left(\left(\frac{z_{\min}}{z} \right)^{\alpha_z} - (1 - \mu) \left(\frac{z_{\min}}{z_m} \right)^{\alpha_z} \right). \quad (29)$$

For $z \geq z_p$, $m(z) = \overline{G}_x^{-1}(r(z))$, so that (29) implies

$$m(z) = x_{\min} \lambda^{-\frac{1}{\alpha_x}} \left(\left(\frac{z_{\min}}{z} \right)^{\alpha_z} - (1 - \mu) \left(\frac{z_{\min}}{z_m} \right)^{\alpha_z} \right)^{-\frac{1}{\alpha_x}} \text{ for } z \in (z_p, z_m).$$

(3) still applies and now gives the differential equation:

$$\left(w'(z) z + \frac{\beta_z}{1 - \beta_z} w(z) \right) = \frac{\beta_z}{1 - \beta_z} x_{\min} \lambda^{\frac{\alpha_x - 1}{\alpha_x}} \left(\left(\frac{z_{\min}}{z} \right)^{\alpha_z} - (1 - \mu) \left(\frac{z_{\min}}{z_m} \right)^{\alpha_z} \right)^{-\frac{1}{\alpha_x}}.$$

Using that $w(z_m) = x_{\min} \left(\frac{z_m}{z_{\min}} \right)^{\frac{\alpha_z}{\alpha_x}}$, the solution to this differential equation is then given by:

$$w(z) = z^{-\frac{\beta_z}{1 - \beta_z}} x_{\min} (z_{\min})^{-\frac{\alpha_z}{\alpha_x}} \left(z_m^{\frac{\alpha_z}{\alpha_x} + \frac{\beta_z}{1 - \beta_z}} - \frac{\beta_z}{1 - \beta_z} \lambda^{\frac{\alpha_x - 1}{\alpha_x}} \int_z^{z_m} \zeta^{-\frac{1 - 2\beta_z}{1 - \beta_z}} (\zeta^{-\alpha_z} - (1 - \mu) z_m^{-\alpha_z})^{-\frac{1}{\alpha_x}} d\zeta \right).$$

For this to be an equilibrium, we need to check that $w(z) \geq x_{\min} \left(\frac{z}{z_{\min}} \right)^{\frac{\alpha_z}{\alpha_x}}$, which is the income that a doctor of rank $\overline{G}_z(z)$ would obtain as a consumer. We can rewrite:

$$w(z) - x_{\min} \left(\frac{z}{z_{\min}} \right)^{\frac{\alpha_z}{\alpha_x}} = x_{\min} (z_{\min})^{-\frac{\alpha_z}{\alpha_x}} z^{-\frac{\beta_z}{1 - \beta_z}} T(z)$$

with

$$T(z) \equiv z_m^{\frac{\alpha_z}{\alpha_x} + \frac{\beta_z}{1 - \beta_z}} - z^{\frac{\alpha_z}{\alpha_x} + \frac{\beta_z}{1 - \beta_z}} - \frac{\beta_z}{1 - \beta_z} \lambda^{\frac{\alpha_x - 1}{\alpha_x}} \int_z^{z_m} \zeta^{-\frac{1 - 2\beta_z}{1 - \beta_z}} (\zeta^{-\alpha_z} - (1 - \mu) z_m^{-\alpha_z})^{-\frac{1}{\alpha_x}} d\zeta.$$

We get

$$T'(z) = \left(1 - \left(\frac{z^{-\alpha_z} - (1 - \mu) z_m^{-\alpha_z}}{\mu z^{-\alpha_z}} \right)^{\frac{1}{\alpha_x}} \right) \frac{\beta_z \lambda^{\frac{\alpha_x - 1}{\alpha_x}}}{1 - \beta_z} z^{-\frac{1 - 2\beta_z}{1 - \beta_z}} (z^{-\alpha_z} - (1 - \mu) z_m^{-\alpha_z})^{-\frac{1}{\alpha_x}}.$$

where we used that

$$\frac{\alpha_z}{\alpha_x} \frac{1 - \beta_z}{\beta_z} + 1 = \lambda (\mu \lambda)^{-\frac{1}{\alpha_x}}. \quad (30)$$

Further for $z < z_m$, we get that $z^{-\alpha_z} - (1 - \mu) z_m^{-\alpha_z} > \mu z^{-\alpha_z}$, so that $T'(z) < 0$. Since $T(z_m) = 0$, then we get that $T(z) > 0$ for $z < z_m$, which ensures that $w(z) > x_{\min} \left(\frac{z}{z_{\min}} \right)^{\frac{\alpha_z}{\alpha_x}}$ for $z_p \leq z < z_m$.

Finally, we consider what happens for $z < z_p$. Denote by $d(z)$ the doctor's ability of the

individual of rank $r(z)$, then using (29) we get:

$$d(z) = \lambda^{-\frac{1}{\alpha_z}} \left(z^{-\alpha_z} - (1 - \mu) z_m^{-\alpha_z} \right)^{-\frac{1}{\alpha_z}}. \quad (31)$$

To close the market, it must be that $d(z_{\min}) = z_{\min}$, which implies that

$$z_m = z_{\min} \left(\frac{1 - \mu}{1 - \frac{1}{\lambda}} \right)^{\frac{1}{\alpha_z}}. \quad (32)$$

Therefore $z_m > z_{\min}$ is only possible if $\mu < 1/\lambda$, which corresponds to $\lambda < \frac{\alpha_z}{\alpha_x} \frac{1 - \beta_z}{\beta_z} + 1$ (the opposite from case 1).

Further, by definition again, we must have $d(z_p) = z_m$, so that:

$$z_p = \frac{z_m}{\left(1 + \frac{1}{\lambda} - \mu\right)^{\frac{1}{\alpha_z}}} = z_{\min} \left(\frac{1 - \mu}{\left(1 - \frac{1}{\lambda}\right) \left(1 + \frac{1}{\lambda} - \mu\right)} \right)^{\frac{1}{\alpha_z}}. \quad (33)$$

It is direct to verify that for $\mu < 1/\lambda$, $z_{\min} < z_p < z_m$.

Now the patient of the doctor of quality z will have an income given by $w(d(z))$. Therefore (3) gives that for $z \leq z_p$, $w(z)$ must satisfy:

$$w'(z) z = \frac{\beta_z}{1 - \beta_z} (\lambda w(d(z)) - w(z)).$$

Multiply this equation by $z^{\frac{\beta_z}{1 - \beta_z} - 1}$ and integrate over (z, z_p) to obtain that the solution must satisfy:

$$w(z) = \left(w(z_p) z_p^{\frac{\beta_z}{1 - \beta_z}} - \int_z^{z_p} \frac{\beta_z}{1 - \beta_z} \zeta^{\frac{2\beta_z - 1}{1 - \beta_z}} \lambda w(d(\zeta)) d\zeta \right) z^{-\frac{\beta_z}{1 - \beta_z}} \text{ for } z \leq z_p.$$

Once again, we need to verify that $w(z) \geq x_{\min} \left(\frac{z}{z_{\min}} \right)^{\frac{\alpha_z}{\alpha_x}}$ for $z < z_p$. Taking the difference

we can write:

$$\begin{aligned}
& w(z) - x_{\min} \left(\frac{z}{z_{\min}} \right)^{\frac{\alpha_z}{\alpha_x}} \\
&= \left[\left(w(z_p) - x_{\min} \left(\frac{z_p}{z_{\min}} \right)^{\frac{\alpha_z}{\alpha_x}} \right) z_p^{\frac{\beta_z}{1-\beta_z}} + \frac{x_{\min}}{z_{\min}^{\frac{\alpha_z}{\alpha_x}}} \left(z_p^{\frac{\alpha_z}{\alpha_x} + \frac{\beta_z}{1-\beta_z}} - z^{\frac{\alpha_z}{\alpha_x} + \frac{\beta_z}{1-\beta_z}} \right) \right. \\
&\quad \left. - \int_z^{z_p} \frac{\lambda \beta_z \zeta^{\frac{2\beta_z-1}{1-\beta_z}}}{1-\beta_z} w(d(\zeta)) d\zeta \right] z^{-\frac{\beta_z}{1-\beta_z}}.
\end{aligned}$$

We already know that $w(z_p) > x_{\min} z_{\min}^{-\frac{\alpha_z}{\alpha_x}} z_p^{\frac{\alpha_z}{\alpha_x}}$. Moreover for $\zeta \in (z, z_p)$, $d(\zeta) < z_m$, since $w(z)$ is increasing we get

$$w(d(\zeta)) \leq w(z_m) = x_{\min} \left(\frac{z_m}{z_{\min}} \right)^{\frac{\alpha_z}{\alpha_x}}.$$

Therefore, we get:

$$w(z) - x_{\min} \left(\frac{z}{z_{\min}} \right)^{\frac{\alpha_z}{\alpha_x}} > \frac{x_{\min} z^{-\frac{\beta_z}{1-\beta_z}}}{z_{\min}^{\frac{\alpha_z}{\alpha_x}}} T_2(z).$$

with

$$T_2(z) = \left(z_p^{\frac{\alpha_z}{\alpha_x} + \frac{\beta_z}{1-\beta_z}} - z^{\frac{\alpha_z}{\alpha_x} + \frac{\beta_z}{1-\beta_z}} \right) - \lambda z_m^{\frac{\alpha_z}{\alpha_x}} \left(z_p^{\frac{\beta_z}{1-\beta_z}} - z^{\frac{\beta_z}{1-\beta_z}} \right).$$

Differentiating, we get:

$$T_2'(z) = \left(\lambda \frac{\beta_z}{1-\beta_z} z_m^{\frac{\alpha_z}{\alpha_x}} - \left(\frac{\alpha_z}{\alpha_x} + \frac{\beta_z}{1-\beta_z} \right) z^{\frac{\alpha_z}{\alpha_x}} \right) z^{\frac{\beta_z}{1-\beta_z}-1}.$$

Therefore $T_2'(z)$ has the sign of $\lambda \frac{\beta_z}{1-\beta_z} z_m^{\frac{\alpha_z}{\alpha_x}} - \left(\frac{\alpha_z}{\alpha_x} + \frac{\beta_z}{1-\beta_z} \right) z^{\frac{\alpha_z}{\alpha_x}}$, which is more likely to be negative for a higher z and can change sign at most once on (z_{\min}, z_p) . Using (33) and (30) we get that

$$T_2'(z_p) = \frac{\beta_z \lambda^{\frac{\alpha_x-1}{\alpha_x}} \mu^{-\frac{1}{\alpha_x}}}{1-\beta_z} z^{\frac{\beta_z}{1-\beta_z}-1} z_p^{\frac{\alpha_z}{\alpha_x}} \left(\left(\left(1 + \frac{1}{\lambda} - \mu \right) \lambda \mu \right)^{\frac{1}{\alpha_x}} - 1 \right).$$

Note that $(1 + \frac{1}{\lambda} - \mu) \lambda \mu = 1 - (1 - \mu)(1 - \lambda \mu)$, since $\lambda \mu < 1$ and $\lambda > 1$ (which implies $\mu < 1$), then we get $(1 + \frac{1}{\lambda} - \mu) \lambda \mu < 1$. Therefore $T_2'(z_p) < 0$, so that over (z_{\min}, z_p) either T_2 is everywhere decreasing or T_2 is initially increasing and afterwards decreasing. In the former case since $T_2(z_p) > 0$, we directly get that $T_2(z) > 0$ for $z \in (z_{\min}, z_p)$. In the latter

case, a necessary and sufficient condition to get $T_2(z) > 0$ over the intervall (z_{\min}, z_p) is that $T_2(z_{\min}) > 0$.

Using (32) and (33), we now compute

$$T_2(z_{\min}) = z_{\min}^{\frac{\alpha_z}{\alpha_x} + \frac{\beta_z}{1-\beta_z}} \cdot \left[\lambda \left(\frac{1-\mu}{1-\frac{1}{\lambda}} \right)^{\frac{1}{\alpha_x}} - 1 - \left(\lambda - \left(\frac{1}{1+\frac{1}{\lambda}-\mu} \right)^{\frac{1}{\alpha_x}} \right) \left(\frac{1-\mu}{1-\frac{1}{\lambda}} \right)^{\frac{1}{\alpha_x}} \left(\frac{1-\mu}{(1-\frac{1}{\lambda})(1+\frac{1}{\lambda}-\mu)} \right)^{\frac{1}{\alpha_z} \frac{\beta_z}{1-\beta_z}} \right]$$

Note that $\lambda - \left(\frac{1}{1+\frac{1}{\lambda}-\mu} \right)^{\frac{1}{\alpha_x}} > 0$ since $\frac{1}{\lambda} > \mu$ and that $\frac{1-\mu}{(1-\frac{1}{\lambda})(1+\frac{1}{\lambda}-\mu)} > 1$ so that $\left(\frac{1-\mu}{(1-\frac{1}{\lambda})(1+\frac{1}{\lambda}-\mu)} \right)^{\frac{1}{\alpha_z} \frac{\beta_z}{1-\beta_z}} > 1$, therefore:

$$\begin{aligned} T_2(z_{\min}) &> z_{\min}^{\frac{\alpha_z}{\alpha_x} + \frac{\beta_z}{1-\beta_z}} \left[\lambda \left(\frac{1-\mu}{1-\frac{1}{\lambda}} \right)^{\frac{1}{\alpha_x}} - 1 - \left(\lambda - \left(\frac{1}{1+\frac{1}{\lambda}-\mu} \right)^{\frac{1}{\alpha_x}} \right) \left(\frac{1-\mu}{1-\frac{1}{\lambda}} \right)^{\frac{1}{\alpha_x}} \right] \\ &\geq z_{\min}^{\frac{\alpha_z}{\alpha_x} + \frac{\beta_z}{1-\beta_z}} \left[\left(\frac{1-\mu}{(1-\frac{1}{\lambda})(1+\frac{1}{\lambda}-\mu)} \right)^{\frac{1}{\alpha_x}} - 1 \right] \\ &> 0, \end{aligned}$$

since $\frac{1-\mu}{(1-\frac{1}{\lambda})(1+\frac{1}{\lambda}-\mu)} > 1$. This guarantees that we always have $T_2(z) > 0$ over (z_{\min}, z_p) , so that we obtain $w(z) > x_{\min} \left(\frac{z}{z_{\min}} \right)^{\frac{\alpha_z}{\alpha_x}}$ for $z \in (z_{\min}, z_m)$, which ensures that we do have an equilibrium: no doctor of rank higher than $\bar{G}_z(z_m)$ would like to switch and be a consumer.

Case 3. We now consider the case where below a certain rank $\bar{G}_z(z_1)$ all individuals choose to be doctors, while above that rank some individuals choose to be consumers.

Consider a $\delta > 0$ and an individual whose ability as a doctor $z \in (z_1, z_1 + \delta)$. Since $\lambda > 1$, labor market clearing imposes that for δ_1 small enough that individual will cure somebody whose rank is above $\bar{G}_z(z_1)$. Therefore the income of the patient is equal to what he would earn as a consumer (since either he is a consumer or must be indifferent between being a doctor himself or a consumer). We can then write labor market clearing as:

$$\left(\frac{x_{\min}}{m(z)} \right)^{\alpha_x} = \lambda \left(\frac{z_{\min}}{z} \right)^{\alpha_z},$$

so that

$$m(z) = x_{\min} \lambda^{-\frac{1}{\alpha_x}} \left(\frac{z}{z_{\min}} \right)^{\frac{\alpha_z}{\alpha_x}}.$$

Using the first order condition (3), we get

$$w'(z)z + \frac{\beta_z}{1-\beta_z}w(z) = \frac{\beta_z}{1-\beta_z}\lambda^{1-\frac{1}{\alpha_x}}x_{\min}\left(\frac{z}{z_{\min}}\right)^{\frac{\alpha_z}{\alpha_x}}.$$

Multiplying on both sides by $z^{\frac{\beta_z}{1-\beta_z}-1}$ and integrate over (z_1, z) to obtain

$$\begin{aligned} \int_{z_1}^z \left(w'(\zeta)\zeta + \frac{\beta_z}{1-\beta_z}w(\zeta) \right) \zeta^{\frac{\beta_z}{1-\beta_z}-1} d\zeta &= \int_{z_1}^z \frac{\beta_z}{1-\beta_z}\lambda^{1-\frac{1}{\alpha_x}}x_{\min}\left(\frac{\zeta}{z_{\min}}\right)^{\frac{\alpha_z}{\alpha_x}} \zeta^{\frac{\beta_z}{1-\beta_z}-1} d\zeta \\ \Rightarrow w(z)z^{\frac{\beta_z}{1-\beta_z}} - w(z_1)z_1^{\frac{\beta_z}{1-\beta_z}} &= Mx_{\min}\left(\left(\frac{z}{z_{\min}}\right)^{\frac{\alpha_z}{\alpha_x}} z^{\frac{\beta_z}{1-\beta_z}} - \left(\frac{z_1}{z_{\min}}\right)^{\frac{\alpha_z}{\alpha_x}} z_1^{\frac{\beta_z}{1-\beta_z}} \right). \end{aligned} \quad (34)$$

where we define $M \equiv \frac{\lambda^{1-\frac{1}{\alpha_x}}}{\frac{\alpha_z}{\alpha_x}\frac{1-\beta_z}{\beta_z}+1}$. Assume that $M < 1$ then we would get $w(z)z^{\frac{\beta_z}{1-\beta_z}} < x_{\min}\left(\frac{z}{z_{\min}}\right)^{\frac{\alpha_z}{\alpha_x}}$, for $z > z_1$, contradicting the fact that this individual chooses to be a doctor. Therefore we must have $M \geq 1$.

Assume now that there is a $\delta > 0$, such that individuals ranked between $\overline{G}_z(z_1)$ and $\overline{G}_z(z_1 - \delta)$ are indifferent between being doctors and being consumer. Consider $z \in (z_1 - \delta, z_1)$, then individual with such ability is indifferent between being a doctor and a consumer so that $w(z) = x_{\min}\left(\frac{z}{z_{\min}}\right)^{\frac{\alpha_z}{\alpha_x}}$. Health care market clearing can be written as:

$$\left(\frac{x_{\min}}{m(z)}\right)^{\alpha_x} = \lambda \left(\left(\frac{z_{\min}}{z_1}\right)^{\alpha_z} + \int_z^{z_1} \alpha_z \mu(\zeta) \zeta^{-\alpha_z-1} z_{\min}^{\alpha_z} d\zeta \right).$$

This implies that

$$m(z) = x_{\min} \lambda^{-\frac{1}{\alpha_x}} \left(\left(\frac{z_{\min}}{z_d}\right)^{\alpha_z} + \int_z^{z_1} \alpha_z \mu(\zeta) \zeta^{-\alpha_z-1} z_{\min}^{\alpha_z} d\zeta \right)^{-\frac{1}{\alpha_x}}.$$

Plugging in this expression in the first order condition (3) which still holds we obtain:

$$\left(\frac{1}{z_1}\right)^{\alpha_z} + \int_z^{z_1} \alpha_z \mu(\zeta) \zeta^{-\alpha_z-1} d\zeta = \frac{\lambda^{\alpha_x-1}}{\left(\frac{\alpha_z}{\alpha_x}\frac{1-\beta_z}{\beta_z} + 1\right)^{\alpha_x}} z^{-\alpha_z}.$$

Differentiating with respect to z , one gets that $\mu(z) = M^{\alpha_x}$, since $M \geq 1$, then $\mu(z) \geq 1$. We assumed that $\mu(z) < 1$, therefore there is a contradiction: individuals cannot be indifferent between being consumers and doctors. Instead all individuals choose to be

consumers.

Therefore the equilibrium must be such that all individuals rank above $\bar{G}_z(z_1)$ are consumers and all ranked below are doctors. Using market clearing for the whole population, we get

$$1 = \lambda (z_1/z_{\min})^{\alpha z} \Rightarrow z_1 = \lambda^{\frac{1}{\alpha z}} z_{\min}.$$

More generally market clearing above z_1 implies that for $z > z_1$, $r(z) = \lambda \left(\frac{z_{\min}}{z}\right)^{\alpha z}$, where, as before, $r(z)$ denotes the rank of the patient of a doctor of quality z . Define z_2 such that $r(z_2) = z_1$, so that $z_2 = \lambda^{\frac{1}{\alpha z}} z_1$, that is z_2 is the ability of the doctor who cures patients of doctor's ability z_1 . Then all doctors with ability in (z_1, z_2) will cure consumers, while all doctors with ability higher than z_2 will cure doctors (with ability $d(z) = \lambda^{-\frac{1}{\alpha z}} z$).

Equation (34) applies on (z_1, z_2) , so that one gets

$$w(z) \geq x_{\min} \left(\frac{z}{z_{\min}} \right)^{\frac{\alpha z}{\alpha x}}.$$

Doctors do not have an incentive to deviate and become consumers.

For $z \geq z_2$, then we get that doctors cure other doctors with ability $d(z) = \lambda^{-\frac{1}{\alpha z}} z$, (3) then leads to:

$$w'(z)z + \frac{\beta_z}{1-\beta_z} w(z) = \lambda w \left(\lambda^{-\frac{1}{\alpha z}} z \right).$$

Define $z_i = \lambda^{\frac{1}{\alpha z}} z_{\min}$, and assume that over (z_{i-1}, z_i) , we have that $w(z) \geq x_{\min} \left(\frac{z}{z_{\min}} \right)^{\frac{\alpha z}{\alpha x}}$ (which is true for $i \in \{1, 2\}$) Then for $z \in (z_i, z_{i+1})$:

$$\int_{z_i}^z \left(w'(\zeta) \zeta + \frac{\beta_z}{1-\beta_z} w(\zeta) \right) \zeta^{\frac{\beta_z}{1-\beta_z}-1} d\zeta = \lambda \int_{z_i}^z w \left(\lambda^{-\frac{1}{\alpha z}} \zeta \right) \zeta^{\frac{\beta_z}{1-\beta_z}-1} d\zeta.$$

Using that $w \left(\lambda^{-\frac{1}{\alpha z}} \zeta \right) \geq x_{\min} \left(\frac{\lambda^{-\frac{1}{\alpha z}} \zeta}{z_{\min}} \right)^{\frac{\alpha z}{\alpha x}}$ since $\lambda^{-\frac{1}{\alpha z}} \zeta \in (z_{i-1}, z_i)$, one gets:

$$w(z) z^{\frac{\beta_z}{1-\beta_z}} - w(z_i) z_i^{\frac{\beta_z}{1-\beta_z}} \geq M x_{\min} \left(\left(\frac{z}{z_{\min}} \right)^{\frac{\alpha z}{\alpha x}} z^{\frac{\beta_z}{1-\beta_z}} - \left(\frac{z_i}{z_{\min}} \right)^{\frac{\alpha z}{\alpha x}} z_i^{\frac{\beta_z}{1-\beta_z}} \right),$$

as $M \geq 1$, then

$$w(z) z^{\frac{\beta_z}{1-\beta_z}} \geq x_{\min} \left(\frac{z}{z_{\min}} \right)^{\frac{\alpha z}{\alpha x}} z^{\frac{\beta_z}{1-\beta_z}} + \left(w(z_i) - \left(\frac{z_i}{z_{\min}} \right)^{\frac{\alpha z}{\alpha x}} \right) z_i^{\frac{\beta_z}{1-\beta_z}},$$

$$\Rightarrow w(z) = x_{\min} \left(\frac{z}{z_{\min}} \right)^{\frac{\alpha_z}{\alpha_x}}.$$

Therefore by recursivity, we get that for all $z \geq z_1$, doctors do not have an incentive to become consumers, which ensures that this is indeed an equilibrium.

Summary. Consider $M = \frac{\lambda^{1-\frac{1}{\alpha_x}}}{\frac{\alpha_z}{\alpha_x} \frac{1-\beta_z}{\beta_z} + 1}$ We have three cases:

- If $M \geq 1$, then individuals rank below $\bar{G}_z \left(\lambda^{\frac{1}{\alpha_z}} z_1 \right)$ are all doctors those above are all consumers.
- If $M < 1$ and $\lambda \geq \frac{\alpha_z}{\alpha_x} \frac{1-\beta_z}{\beta_z} + 1$, a fixed share $\mu = M^{\alpha_x}$ choose to be doctors below a certain rank and all choose to be consumers above that rank.
- If $M < 1$ and $\lambda < \frac{\alpha_z}{\alpha_x} \frac{1-\beta_z}{\beta_z} + 1$. a fixed share $\mu = M^{\alpha_x}$ choose to be doctors below a certain rank and all choose to be doctors above that rank.

B.2.2 Disentangling supply side and demand side effects

To disentangle supply-side and demand side effects in section 2.2.2, we now build a model where doctors have an outside option positively correlated with their ability but where patients are a separate group. Formally, there are two types of agents: a mass 1 of consumers, with income x distributed with the Pareto distribution $P(X > x) = (x_{\min}/x)^{\alpha_x}$ and a mass M of potential doctors. Consumers consume the homogeneous good and health care services according to the utility function 1. Potential doctors only consume the homogeneous good, as in section 2.2.2, they are ranked in descending order of ability and we denote i their rank. Agent i can choose between being a doctor providing health services of quality $z(i)$ and earning $w(z(i))$ or working in the homogeneous good sector earning $y(i)$. y and z are distributed according to the countercumulative distributions:

$$\bar{G}_y(y(i)) = \bar{G}_z(z(i)) = i \text{ with } \bar{G}_y = \left(\frac{y_{\min}}{y} \right)^{\alpha_y} \text{ and } \bar{G}_z = \left(\frac{z_{\min}}{z} \right)^{\alpha_z}.$$

Further $\lambda M > 1$ and $\lambda > 1$ so that everybody can get health services.

Assume that the equilibrium is such that for individuals of a sufficiently high level of ability, some will choose to be doctors and others to work in the homogeneous good sector. That is for i low enough, agents must be indifferent between becoming a doctor or working in the homogeneous good sector, so that we must have $w(z(i)) = y(i)$. Hence, the wage

function must satisfy:

$$w(z) = y_{\min} \left(\frac{z}{z_{\min}} \right)^{\frac{\alpha_z}{\alpha_y}}. \quad (35)$$

Market clearing for health care services above z implies:

$$\left(\frac{x_{\min}}{m(z)} \right)^{\alpha_x} = \lambda M \int_z^{\infty} \mu(\zeta) g_z(\zeta) d\zeta, \quad (36)$$

where $\mu(\zeta)$ denotes the share of potential doctors who decide to work as doctors. Hence:

$$m(z) = x_{\min} \left(\int_z^{\infty} \lambda M \mu(\zeta) g_z(\zeta) d\zeta \right)^{-\frac{1}{\alpha_x}}.$$

Plugging this expression in the first order condition (3) together with (35), we obtain:

$$\int_z^{\infty} \mu(\zeta) g_z(\zeta) d\zeta = \frac{1}{\lambda M} \left(\frac{\frac{\beta_z}{1-\beta_z} \lambda x_{\min}}{\left(\frac{\alpha_z}{\alpha_y} + \frac{\beta_z}{1-\beta_z} \right) y_{\min}} \right)^{\alpha_x} \left(\frac{z}{z_{\min}} \right)^{-\alpha_x \frac{\alpha_z}{\alpha_y}}. \quad (37)$$

Taking the derivative with respect to z , we get:

$$\mu(z) = \frac{\alpha_x}{\alpha_y} \frac{1}{\lambda M} \left(\frac{\frac{\beta_z}{1-\beta_z} \lambda x_{\min}}{\left(\frac{\alpha_z}{\alpha_y} + \frac{\beta_z}{1-\beta_z} \right) y_{\min}} \right)^{\alpha_x} \left(\frac{z}{z_{\min}} \right)^{\alpha_z \left(1 - \frac{\alpha_x}{\alpha_y} \right)}. \quad (38)$$

Since $\mu(z) \in (0, 1)$, this case is only possible if $\alpha_y \leq \alpha_x$, that is consumers income distribution has a fatter tail than the outside option for potential doctors (and $\frac{\alpha_x}{\alpha_y} \frac{1}{\lambda M} \left(\frac{\alpha_y \beta_z \lambda x_{\min}}{(\alpha_z(1-\beta_z) + \beta_z \alpha_y) y_{\min}} \right)^{\alpha_x} \leq 1$ if $\alpha_y = \alpha_x$). We then obtain that doctor's income distribution obeys (for w high enough):

$$\Pr(W > w) = \int_{z_{\min} \left(\frac{w}{y_{\min}} \right)^{\frac{\alpha_y}{\alpha_z}}}^{\infty} \mu(\zeta) \left(\frac{z_{\min}}{\zeta} \right)^{\alpha_z} \frac{d\zeta}{\zeta} = \frac{1}{\lambda M \alpha_z} \left(\frac{\alpha_y \beta_z \lambda x_{\min}}{\alpha_z (1 - \beta_z) + \beta_z \alpha_y} \right)^{\alpha_x} w^{-\alpha_x}.$$

Therefore doctors' income is distributed like the patients' income and not according to doctors' outside option.

With $\alpha_y > \alpha_x$ or $\alpha_y = \alpha_x$ together with $\frac{\alpha_x}{\alpha_y} \frac{1}{\lambda M} \left(\frac{\alpha_y \beta_z \lambda x_{\min}}{(\alpha_z(1-\beta_z) + \beta_z \alpha_y) y_{\min}} \right)^{\alpha_x} > 1$, then above a certain threshold, all potential doctors will choose to be doctors, so that the model behaves like that of section 2.1.

Therefore, in all cases, at the top, income is distributed in a Pareto way with shape parameter α_x . Changes in α_y have no impact on doctors' top income inequality.

B.3 Mobility and Open Economy

B.3.1 Doctors moving

We return again to the baseline model of section 2.1, but we now assume that there are 2 regions, A and B , and that doctors can move across regions.³⁰ But medical services are again non-tradable and patients cannot move.³¹ The two regions are identical except for the ability distribution of consumers, which is Pareto in both but with possibly different means and shape parameters.³² Without loss of generality, we assume that $\alpha_x^A < \alpha_x^B$; that is region A is more unequal than region B .

With no trade in goods between the two regions, we can normalize the price of the homogeneous good to 1 in both. As doctors only consume the homogeneous good, doctors' nominal wages must be equalized in the two regions. As a result the price of health care of quality z must be the same in both regions. From the first order condition on health care consumption, this implies that the matching function is the same: doctors of quality z provide health care to consumers of income $m(z)$ in both regions. Moreover, the least able potential doctor who decides to become a doctor must have the same ability z_c in both regions.

We define by $\varphi(z)$ the *net* share of doctors initially in region B with ability *at least* z who decide to move to region A . Then labor market clearing in region A implies that, for $z \geq z_c$,

$$(x_{\min}^A/m(z))^{\alpha_x^A} = \lambda\mu_d(1 + \varphi(z))(z_{\min}/z)^{\alpha_z}. \quad (39)$$

There are initially $\mu_d(z_{\min}/z)^{\alpha_z}$ doctors with ability at least z in each region and by definition, a share $\varphi(z)$ of those move from region B to region A . Since each doctor can provide services to λ patients, after doctors have relocated the total supply over a quality z in region A is given by the right-hand side of (39). Total demand corresponds to region A patients with an income higher than $m(z)$, of which there are $P(X > m(z))$. The same equation, replacing $\varphi(z)$ by $-\varphi(z)$, holds in region B :

$$(x_{\min}^B/m(z))^{\alpha_x^B} = \lambda\mu_d(1 - \varphi(z))(z_{\min}/z)^{\alpha_z}. \quad (40)$$

Since the two regions are of equal size, total demand for health services must be the

³⁰The results generalize to more than 2 regions.

³¹When doctors are mobile and medical services tradable, the geographic location of agents is undetermined in general, and we would need a full spatial equilibrium model to generate empirical predictions.

³²Our results directly generalize to a case where the two regions do not have the same mass of potential doctors and consumers.

same and on net, no doctors move: $\varphi(z_c) = 0$. On the other hand, most rich patients are in region A (as $\alpha_x^A > \alpha_x^B$). As doctors' incomes increase with the incomes of their patients, nearly all of the most talented doctors will eventually locate in region A : $\lim_{z \rightarrow \infty} \varphi(z) = 1$. We therefore obtain that, in region A , the distribution of doctors' ability after relocation is asymptotically Pareto. So, as in the baseline model, doctors' incomes will be asymptotically Pareto distributed with a shape parameter equal to α_x^A .

In region B , doctors of a given quality level earn the same as in region A . That is, the incomes of doctors initially in region B are still Pareto distributed with coefficient α_x^A . However, after the move, the share of doctors that stay in region B decreases with their quality. Using (39) and (40), we get that $1 - \varphi(z) \propto z^{\alpha_z(1 - \alpha_x^B/\alpha_x^A)}$. Therefore, the *ex post* talent distribution of doctors in region B is still Pareto but now with a coefficient $\alpha'_z = \alpha_z \alpha_x^B / \alpha_x^A$. As in the baseline model, the distribution of income for doctors who stay must be asymptotically Pareto with a shape Parameter α_x^B .³³ We obtain:

Proposition 4. *Once doctors have relocated, the income distribution of doctors in region A is asymptotically Pareto with coefficient α_x^A , and the income distribution of doctors in region B is asymptotically Pareto with coefficient α_x^B .*

Formal proof is in Appendix B.3.2 below.

Consequently, whether doctors can move or not does not alter the *observable* local income distribution, although it does matter considerably for the unobservable local *ability* distribution. Consequently, for our empirical analysis we need not take a stand on whether doctors are mobile. We cannot empirically distinguish between the free-mobility and no-migration cases using data on income inequality.

B.3.2 Proof of Proposition 3

Since $\omega(z)$ is equalized between the two regions, then the threshold z_c of the least able potential doctor must also be the same in the two regions.³⁴ Summing up the market clearing equations (39) and (40), we obtain that as in the baseline model, $z_c = (\lambda \mu_d)^{\frac{1}{\alpha_z}} z_{min}$.

³³To see that there is no contradiction, note that the baseline model predicts that the income of individual z , $w(z) \propto z^{\frac{\alpha'_z}{\alpha_x^B}}$ but $\frac{\alpha'_z}{\alpha_x^B} = \frac{\alpha_z}{\alpha_x^A}$, so we also have $w(z) \propto z^{\frac{\alpha_z}{\alpha_x^A}}$ and doctors do indeed earn the same in both regions.

³⁴Here potential doctors who decide to work in the homogeneous good sector would go to region B since $\alpha_x^A > \alpha_x^B$ implies that $x_{min}^A < x_{min}^B$. This is without consequences: alternatively, we could have assumed that the outside option of doctors is to produce \hat{x} , which is identical between the two regions. In that case potential doctors who work in the homogeneous sector would not move.

Next combining (39) and (40), we get that

$$x_{\min}^A (1 + \varphi(z))^{-\frac{1}{\alpha_x^A}} = x_{\min}^B \left(\frac{z}{z_c}\right)^{\frac{\alpha_z}{\alpha_x^B} - \frac{\alpha_z}{\alpha_x^A}} (1 - \varphi(z))^{-\frac{1}{\alpha_x^B}}. \quad (41)$$

Since $\alpha_x^B > \alpha_x^A$, we get that $\left(\frac{z}{z_c}\right)^{\frac{\alpha_z}{\alpha_x^B} - \frac{\alpha_z}{\alpha_x^A}}$ tends towards 0. As a net share $\varphi(z) \in (-1, 1)$, if $\varphi(z) \rightarrow -1$, then the left-hand side would tend toward infinity and the right-hand side toward 0, which is a contradiction. Therefore $1 + \varphi(z)$ must be bounded below, which ensures that the left-hand side is bounded above 0. If $\varphi(z) \not\rightarrow 1$, then the right-hand side would be asymptotically 0, this is also a contradiction. Therefore asymptotically, we must have that $\varphi(z) \rightarrow 1$: nearly all the best doctors move to the most unequal region.

Plugging (39) in (3), we get that in region A:

$$w'(z)z + \frac{\beta_z}{1 - \beta_z}w(z) = \frac{\beta_z \lambda}{1 - \beta_z} (1 + \varphi(z))^{-\frac{1}{\alpha_x^A}} \left(\frac{z_c}{z}\right)^{-\frac{\alpha_z}{\alpha_x^A}}.$$

Therefore, asymptotically:

$$w(z) \rightarrow \frac{\lambda \beta_z \alpha_x^A 2^{-\frac{1}{\alpha_x^A}}}{\alpha_z (1 - \beta_z) + \beta_z \alpha_x^A} \left(\frac{z}{z_c}\right)^{\frac{\alpha_z}{\alpha_x^A}} \quad (42)$$

Since $\varphi(z) \rightarrow 1$, after the location decision, doctors' talent is asymptotically distributed with Pareto coefficient α_z in region A: for z high enough, there are $2\mu_d (z_{\min}/z)^{\alpha_z}$ doctors eventually located in region A. We then directly get that doctor's income distribution is asymptotically Pareto distributed with coefficient α_x^A .

From (41), we get that:

$$\begin{aligned} 1 - \varphi(z) &= (x_{\min}^B/x_{\min}^A)^{\alpha_x^B} (1 + \varphi(z))^{\alpha_x^B/\alpha_x^A} (z/z_c)^{\alpha_z(1-\alpha_x^B/\alpha_x^A)} \\ &\rightarrow 2^{\alpha_x^B/\alpha_x^A} (x_{\min}^B/x_{\min}^A)^{\alpha_x^B} (z/z_c)^{\alpha_z(1-\alpha_x^B/\alpha_x^A)}. \end{aligned} \quad (43)$$

Then we can write that in region B, the probability that a doctor earns at least \tilde{w} is given by:

$$P_{doc}^B(W > \tilde{w}) = \frac{\mu_d P(Z > w^{-1}(\tilde{w})) (1 - \varphi(w^{-1}(\tilde{w})))}{\mu_d P(Z > z_c)},$$

where w above denotes the wage function. Indeed, there are originally $\mu_d P(Z > w^{-1}(\tilde{w}))$ doctors present in region B with a talent sufficient to earn \tilde{w} . Out of these doctors, $1 - \varphi(w^{-1}(\tilde{w}))$ stay in region B. Moreover, the total mass of active doctors in region B is given

by $\mu_d P(Z > z_c)$, since overall there is no net movement of actual doctors. Using (42) we get that,

$$w^{-1}(\tilde{w}) \rightarrow z_c \left(\frac{\tilde{w}^{\alpha_z} (1 - \beta_z) + \beta_z \alpha_x^A}{\lambda \beta_z \alpha_x^A} 2^{\frac{1}{\alpha_x^A}} \right)^{\frac{\alpha_x^A}{\alpha_z}}.$$

Using this expression and (43) we get that:

$$P_{doc}^B(W > \tilde{w}) = \left(\frac{z_c}{w^{-1}(\tilde{w})} \right)^{\alpha_z} (1 - \varphi(w^{-1}(\tilde{w}))) \rightarrow \left(\frac{x_{\min}^B}{x_{\min}^A} \frac{\lambda \beta_z \alpha_x^A}{\alpha_z (1 - \beta_z) + \beta_z \alpha_x^A} \frac{1}{\tilde{w}} \right)^{\alpha_x^B}.$$

This establishes Proposition 3/4.

B.4 Utility Function and Ability Distribution

B.4.1 Doctors consume medical services and ability distribution is only Pareto distributed in the tail

We now alter the model so there is a mass 1 of agents, of which a fraction μ_d are potential doctors. The technology for health services is the same as before (and we now assume that $\lambda > 1/\mu_d$). Agents not working as doctors produce a composite good which we take as the numeraire. Unlike in the baseline model, all agents have the same utility function (1).

The equilibrium results in a wage distribution. We assume that this distribution and also the distributions of skills for potential doctors are asymptotically Pareto. Therefore we can write

$$P_x(X > x) = \bar{G}_x(\bar{x}) \bar{G}_{x,\bar{x}}(x),$$

where $\bar{G}_{x,\bar{x}}(x)$ is the conditional counter-cumulative distribution above \bar{x} and $\bar{G}_x(\bar{x})$ is the unconditional counter-cumulative distribution, and for \bar{x} large enough we have

$$\bar{G}_x(x, \bar{x}) \approx (\bar{x}/x)^{\alpha_x},$$

with $\alpha_x > 1$. The same holds for doctors' talents z (moreover, potential doctors can work as consumers with the lowest productivity x_{\min} as an alternative).

As before, solving for the consumer problem leads to the differential equation (3). Furthermore since health care services are not divisible, the equilibrium also features assortative matching and we still denote the matching function $m(z)$. Market clearing at every z can still be written as (4). The least able potential doctor who actually works as a doctor will have ability $z_c = \bar{G}_z^{-1}(1/\lambda\mu_d)$. Therefore z_c is independent of α_x . As a result, (4) implies

that $m(z)$ is defined by $m(z) = \overline{G}_x^{-1}(\overline{G}_{z,z_c}(z))$.

For z above some threshold, \bar{z} , both doctors' talents and incomes are approximately Pareto distributed, which allows us to rewrite the previous equation as:

$$\overline{G}_x(m(\bar{z})) (m(\bar{z}) / (m(z)))^{\alpha_x} \approx \overline{G}_{\bar{z},z_c}(\bar{z}) (\bar{z}/z)^{\alpha_z},$$

which gives

$$m(z) \approx Bz^{\frac{\alpha_z}{\alpha_x}} \text{ with } B = m(\bar{z}) \left(\frac{\overline{G}_x(m(\bar{z}))}{\overline{G}_{\bar{z},z_c}(\bar{z}) \bar{z}^{\alpha_z}} \right)^{\frac{1}{\alpha_x}}.$$

Plugging this in (3) we can rewrite the differential equation as:

$$w'(z)z + \frac{\beta_z}{1-\beta_z}w(z) \approx \frac{\beta_z}{1-\beta_z}\lambda Bz^{\frac{\alpha_z}{\alpha_x}}.$$

Therefore for z large enough, we must have (see Appendix B.4.3 for a derivation):

$$w(z) \approx \frac{\beta_z \alpha_x}{\alpha_z (1-\beta_z) + \beta_z \alpha_x} \lambda B z^{\frac{\alpha_z}{\alpha_x}}. \quad (44)$$

From this we get (as above) that for \bar{w}_d large enough, doctors' income is distributed according to

$$P(W_d > w_d | w_d > \bar{w}_d) \approx \left(\frac{\bar{w}_d}{w_d} \right)^{\alpha_x}. \quad (45)$$

That is, doctors' income follows a Pareto distribution with shape parameter α_x . Proposition 1 still applies: a decrease in α_x will directly translate into an increase in top income inequality among doctors.

B.4.2 The role of the Cobb-Douglas utility function

We keep the same model as just introduced, but we replace the utility function of equation (1) with:

$$u(z, c) = \left(\beta_z z^{\frac{\varepsilon-1}{\varepsilon}} + \beta_c c^{\frac{\varepsilon-1}{\varepsilon}} \right)^{\frac{\varepsilon}{\varepsilon-1}}, \quad (46)$$

with $\varepsilon \neq 1$. As before, the first order conditions gives the differential equation:

$$\partial u / \partial z = \omega'(z) \partial u / \partial c. \quad (47)$$

Since CES exhibits positive cross-partial derivatives, we know that the equilibrium features positive assortative matching. Therefore, with income and ability asymptotically Pareto, the matching function still obeys (5). Using (46), combining (47) and (5), and using that $w(z) = \lambda\omega(z)$, we find that for high levels of z the wage function obeys a differential equation given by

$$w'(z) \approx \lambda \frac{\varepsilon-1}{\varepsilon} \frac{\beta_z}{\beta_c} z^{-\frac{1}{\varepsilon}} \left(\lambda B z^{\frac{\alpha_z}{\alpha_x}} - w(z) \right)^{\frac{1}{\varepsilon}}. \quad (48)$$

We solve this differential equation in Appendix B.4.4 and we prove:

Proposition 5. *i) Assume that $\varepsilon > 1$. Then for $\alpha_x \geq \alpha_z$, wages of doctors are asymptotically Pareto distributed with exponential parameter $\alpha_w = \alpha_x$. For $\alpha_x < \alpha_z$, wages of doctors are asymptotically Pareto distributed with $\alpha_w = \frac{\alpha_z}{\left(\frac{\alpha_z}{\alpha_x} - 1\right)^{\frac{1}{\varepsilon} + 1}}$.
ii) Assume that $\varepsilon < 1$. Then for $\alpha_x > \frac{\alpha_z}{1-\varepsilon}$, wages of doctors are bounded. For $\alpha_x = \frac{\alpha_z}{1-\varepsilon}$, wages of doctors are asymptotically exponentially distributed. For $\alpha_z < \alpha_x < \frac{\alpha_z}{1-\varepsilon}$, they are asymptotically distributed with $\alpha_w = \frac{\alpha_z}{\left(\frac{\alpha_z}{\alpha_x} - 1\right)^{\frac{1}{\varepsilon} + 1}}$. For $\alpha_x \leq \alpha_z$, they are asymptotically Pareto distributed with $\alpha_w = \alpha_x$.*

Therefore, when doctors' income distribution is Pareto, we still obtain that a reduction in α_x leads to a reduction in α_w (that is an increase in general top income inequality increases top income inequality among doctors), although the elasticity may now be lower than 1. (It cannot be asymptotically above 1, since high-paying consumers would then spend more than their income on medical services.) Further, a decrease in α_x also reduces the size of the parameter space for which doctors' wage distribution is bounded (a situation where top income inequality for doctors is very low).

To intuitively understand the results of Proposition 5, consider first the case where $\alpha_z > \alpha_x$. That is, the ability distribution of consumers has a fatter tail than that of doctors, implying a shortage of doctors at the top. This must mean a convex pricing schedule for medical services. If $\varepsilon > 1$, health services and the homogeneous good are substitutes, so the expenditure share on health services declines with income. As a result, $w(z)$ cannot grow as fast as the income of the consumer who buys the services of doctor z , namely $m(z)$, which grows as $z^{\frac{\alpha_z}{\alpha_x}}$. This implies less income inequality among the top doctors than among the top consumers (a higher Pareto exponential parameter). On the other hand, if $\varepsilon < 1$ then richer consumers are forced to spend an increasing amount—eventually all their resources—on health services. So $m(z)$ and $w(z)$ grow at the same rate, and doctors' income is Pareto distributed with coefficient α_x . The reverse holds when doctors are relatively abundant at the top (i.e. when $\alpha_z < \alpha_x$), except that with $\varepsilon < 1$, doctors' income can even be bounded.

B.4.3 Deriving (44)

Using that both doctors talents and income are approximately Pareto distributed, we can rewrite (3) as:

$$(m(\bar{z}) / (m(z)))^{\alpha_x} = \frac{\bar{G}_{\bar{z}, z_c}(\bar{z})}{G_x(m(\bar{z}))} \left((\bar{z}/z)^{\alpha_z} + o\left(\left(\frac{\bar{z}}{z}\right)^{\alpha_z}\right) \right) - o\left(\left(\frac{m(\bar{z})}{m(z)}\right)^{\alpha_x}\right).$$

From this we get that $m(z)$ is of the order of $z^{\frac{\alpha_z}{\alpha_x}}$ and therefore

$$m(z) = Bz^{\frac{\alpha_z}{\alpha_x}} + o\left(z^{\frac{\alpha_z}{\alpha_x}}\right)$$

with B defined as in the text. We can then rewrite (3) as

$$w'(z)z = \frac{\beta_z}{1 - \beta_z} \left(\lambda Bz^{\frac{\alpha_z}{\alpha_x}} - w(z) \right) + o\left(z^{\frac{\alpha_z}{\alpha_x}}\right). \quad (49)$$

We then define $\bar{w}(z) \equiv \frac{\beta_z \alpha_x}{\alpha_z(1-\beta_z) + \beta_z \alpha_x} \lambda Bz^{\frac{\alpha_z}{\alpha_x}}$ which is a solution to the differential equation without the negligible term, and $\tilde{w}(z) \equiv w(z) - \bar{w}(z)$, which must satisfy

$$\tilde{w}'(z)z = -\frac{\beta_z}{1 - \beta_z} \tilde{w}(z) + o\left(z^{\frac{\alpha_z}{\alpha_x}}\right).$$

This gives

$$\tilde{w}'(z)z^{\frac{\beta_z}{1-\beta_z}} + \frac{\beta_z}{1 - \beta_z} \tilde{w}(z)z^{\frac{\beta_z}{1-\beta_z}-1} = o\left(z^{\frac{\alpha_z}{\alpha_x}} z^{\frac{\beta_z}{1-\beta_z}-1}\right)$$

Integrating we obtain:

$$\tilde{w}(z) = Kz^{-\frac{\beta_z}{1-\beta_z}} + o\left(z^{\frac{\alpha_z}{\alpha_x}}\right)$$

for some constant K , therefore $\tilde{w}(z)$ is negligible in front of $\bar{w}(z)$.

B.4.4 Proof of Proposition 5

We rewrite (48) more precisely as:

$$w'(z) = \lambda^{\frac{\varepsilon-1}{\varepsilon}} \frac{\beta_z}{\beta_c} z^{-\frac{1}{\varepsilon}} \left(\lambda Bz^{\frac{\alpha_z}{\alpha_x}} - w(z) \right)^{\frac{1}{\varepsilon}} + o\left(\frac{\lambda Bz^{\frac{\alpha_z}{\alpha_x}} - w(z)}{z}\right)^{\frac{1}{\varepsilon}}. \quad (50)$$

Since consumption of the homogeneous good must remain positive then $\lim \lambda B z^{\frac{\alpha_z}{\alpha_x}} - w(z) \geq 0$, which means that $w(z)$ cannot grow faster than $z^{\frac{\alpha_z}{\alpha_x}}$. We can then distinguish 2 cases: $w(z) = o\left(z^{\frac{\alpha_z}{\alpha_x}}\right)$ and $w(z) \propto z^{\frac{\alpha_z}{\alpha_x}}$.

Case with $w(z) = o\left(z^{\frac{\alpha_z}{\alpha_x}}\right)$. Then for z high enough, one obtains that

$$w'(z) = \lambda \frac{\beta_z}{\beta_c} B^{\frac{1}{\varepsilon}} z^{\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon}} + o\left(z^{\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon}}\right). \quad (51)$$

Integrating, we obtain that for $\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon} \neq -1$

$$w(z) = K + \lambda \frac{\beta_z}{\beta_c} \frac{B^{\frac{1}{\varepsilon}}}{\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon} + 1} z^{\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon} + 1} + o\left(z^{\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon} + 1}\right),$$

where K is a constant. Note that to be consistent, we must have $\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon} + 1 < \frac{\alpha_z}{\alpha_x}$, that is $(\alpha_z - \alpha_x)(\varepsilon - 1) > 0$: this case is ruled out if $\alpha_z \geq \alpha_x$ and $\varepsilon < 1$ or if $\alpha_z \leq \alpha_x$ and $\varepsilon > 1$.

If $\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon} + 1 < 0$ then $w(z)$ is bounded by K .

If $\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon} + 1 > 0$, then we get that

$$w(z) = f^w(z) \equiv \lambda \frac{\beta_z}{\beta_c} \frac{B^{\frac{1}{\varepsilon}}}{\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon} + 1} z^{\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon} + 1} + o\left(z^{\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon} + 1}\right),$$

where the notation f^w is introduced to help notation. Therefore one gets, for \bar{w} large:

$$\Pr(W > w) = \Pr(Z > (f^w)^{-1}(w)) = \bar{G}_w(\bar{w}) \left(\frac{\bar{w}}{w}\right)^{\frac{\alpha_z}{\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon} + 1}} + o\left(w^{-\frac{\alpha_z}{\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon} + 1}}\right),$$

so that w is Pareto distributed asymptotically with a coefficient $\alpha_w = \frac{\alpha_z}{\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon} + 1}$, which is increasing in α_x (and we have $\alpha_w > \alpha_x$).

If $\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon} + 1 = 0$, then $\alpha_z = \alpha_x(1 - \varepsilon)$, and integrating (51), one obtains

$$w(z) = f^w(z) \equiv \lambda \frac{\beta_z}{\beta_c} B^{\frac{1}{\varepsilon}} \ln z + o(\ln z).$$

Therefore

$$\begin{aligned}\Pr(W > w) &= \Pr\left(Z > \left(\exp\left(\frac{\beta_c}{\lambda\beta_z B^{\frac{1}{\varepsilon}}} w\right) + o(\exp(w))\right)\right) \\ &= \bar{G}_{z,z_c}(\bar{z}) \bar{z}^{\alpha_z} \exp\left(-\frac{\alpha_z \beta_c}{\lambda\beta_z B^{\frac{1}{\varepsilon}}} w\right) + o(\exp(-\alpha_z w))\end{aligned}$$

In that case, w is distributed exponentially.

Case where $w(z) \propto z^{\frac{\alpha_z}{\alpha_x}}$. That is we assume that

$$w(z) = Az^{\frac{\alpha_z}{\alpha_x}} + o\left(z^{\frac{\alpha_z}{\alpha_x}}\right) \quad (52)$$

for some constant $A > 0$. Then, we have that

$$\begin{aligned}\Pr(W > w) &= \Pr\left(Z > \left(\left(\frac{w}{A}\right)^{\frac{\alpha_x}{\alpha_z}} + o(w)^{\frac{\alpha_x}{\alpha_z}}\right)\right) \\ &= \bar{G}_w(\bar{w}) \left(\frac{\bar{w}}{w}\right)^{\alpha_x} + o(w)^{\frac{\alpha_x}{\alpha_z}}\end{aligned}$$

That is w is Pareto distributed with coefficient α_x .

Plugging (52) in (50), we get:

$$A \frac{\alpha_z}{\alpha_x} z^{\frac{\alpha_z}{\alpha_x} - 1} + o\left(z^{\frac{\alpha_z}{\alpha_x} - 1}\right) = \lambda \frac{\varepsilon - 1}{\varepsilon} \frac{\beta_z}{\beta_c} (\lambda B - A)^{\frac{1}{\varepsilon}} z^{\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon}} + o\left((\lambda B - A)^{\frac{1}{\varepsilon}} z^{\left(\frac{\alpha_z}{\alpha_x} - 1\right)\frac{1}{\varepsilon}}\right). \quad (53)$$

First, if $\alpha_z = \alpha_x$, then we obtain $A = \lambda \frac{\varepsilon - 1}{\varepsilon} \frac{\beta_z}{\beta_c} (\lambda B - A)^{\frac{1}{\varepsilon}}$.

Consider now that $\alpha_z \neq \alpha_x$. If $\lambda B \neq A$ then (53) is impossible when $\varepsilon \neq 1$, therefore we must have that $\lambda B = A$. This equation then requires that

$$\frac{\alpha_z}{\alpha_x} - 1 < \left(\frac{\alpha_z}{\alpha_x} - 1\right) \frac{1}{\varepsilon} \Leftrightarrow (\alpha_z - \alpha_x)(\varepsilon - 1) < 0.$$

In fact, for $(\alpha_z - \alpha_x)(\varepsilon - 1) < 0$, one gets that

$$w(z) = \lambda B z^{\frac{\alpha_z}{\alpha_x}} - \lambda \left(B \frac{\alpha_z \beta_c}{\alpha_x \beta_z}\right)^{\varepsilon} z^{\varepsilon \left(\frac{\alpha_z}{\alpha_x} - 1\right) + 1} + o\left(z^{\varepsilon \left(\frac{\alpha_z}{\alpha_x} - 1\right) + 1}\right)$$

satisfies (50) provided that the function $o\left(z^{\varepsilon \left(\frac{\alpha_z}{\alpha_x} - 1\right) + 1}\right)$ solves the appropriate differential equation.

Collecting the different cases together gives Proposition 5.

C Data appendix

C.1 Details of data construction

Sample Selection

1. Currently, everyone with non-missing and positive income is included. This includes all ages, and those with missing occupation variables.

Weighting

1. All constructed variables (inequality measures and shares) use census weights. These are based demographic characteristics.

Independent Variable

1. The current code calculates a year-region cutoff, excluding the occupation of interest, corresponding to the 90th percentile of *uncensored* observations.
2. It treats all people above the cutoff as a homogenous occupations: thus capturing within-occupation and across-occupation inequality.

Instrument Occupations

1. The current code takes the top 20 occupations specific to each region in 1980, and uses these for all years' instrument.
2. But, the *national* cutoff is used in each region when defining the top occupations.
3. The current code uses just uncensored observations to calculate the local rankings. This places greater weight on high-earning, but not *very* higher earning occupations.

Instrument Shares

1. Currently, shares are calculated using just *uncensored* observations, corresponding to the rankings above.

Instrument Inequality Measure

1. The current code calculates occupation inequality by taking the nation-wide measure (excluding the local LMA). This captures changes in spatial inequality.

Regressions

1. Regression Weight: Regressions are weighted by the number of outcome occupation contemporary observations above the cutoff in that region.
2. Regression Inclusion: Only LMAs that have at least 8 outcome occupation observations above the cutoff in 1980 are included.
3. We estimate the regressions with the Stata add-on command `xtivreg2`, `fe`.

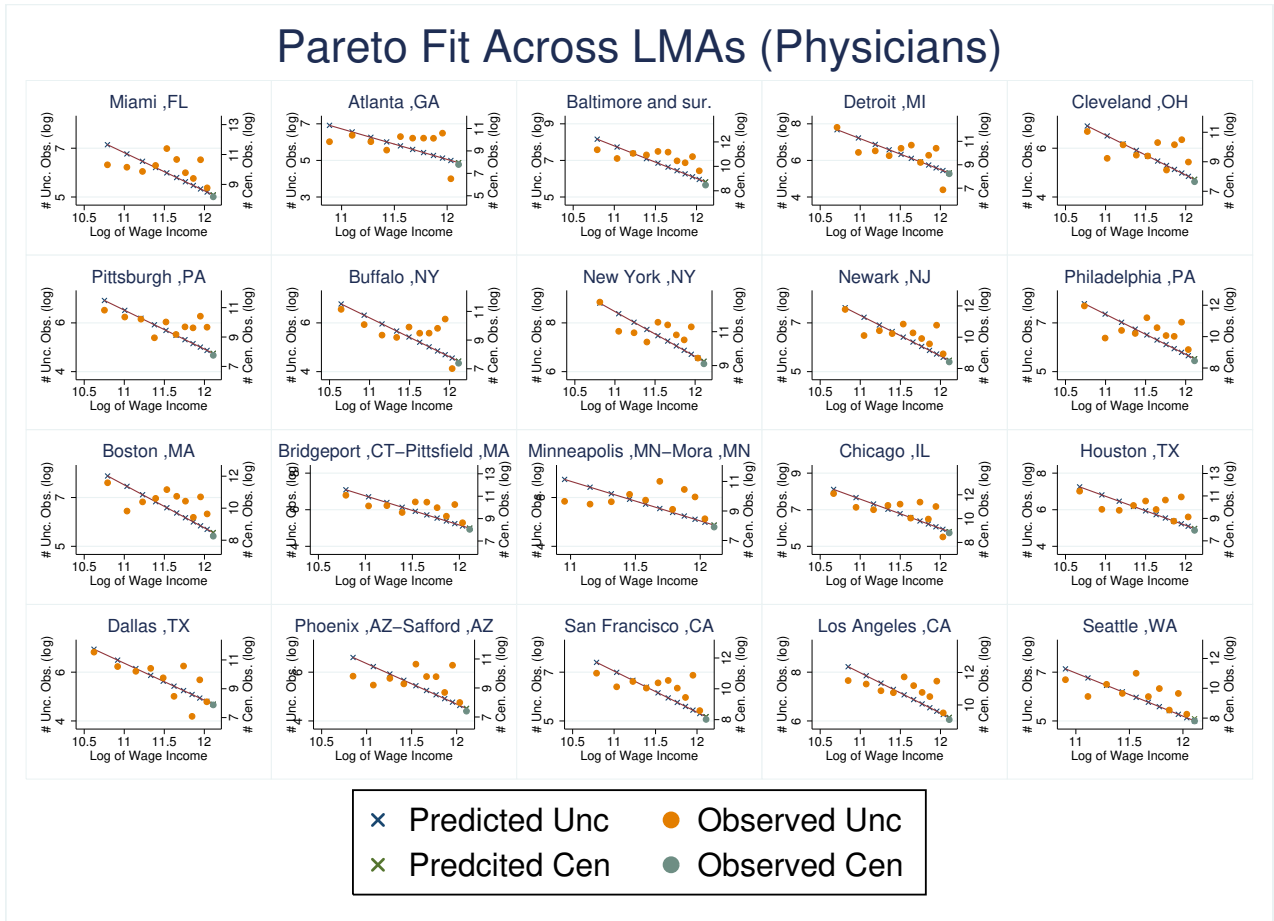
C.2 Statistical method for determining Pareto cutoff

D Pareto Fit and Tables for Top Occupations

Table D.1 gives the change in α for the top occupations. The top occupations for 1980 and 2014 are given in Table D.2.

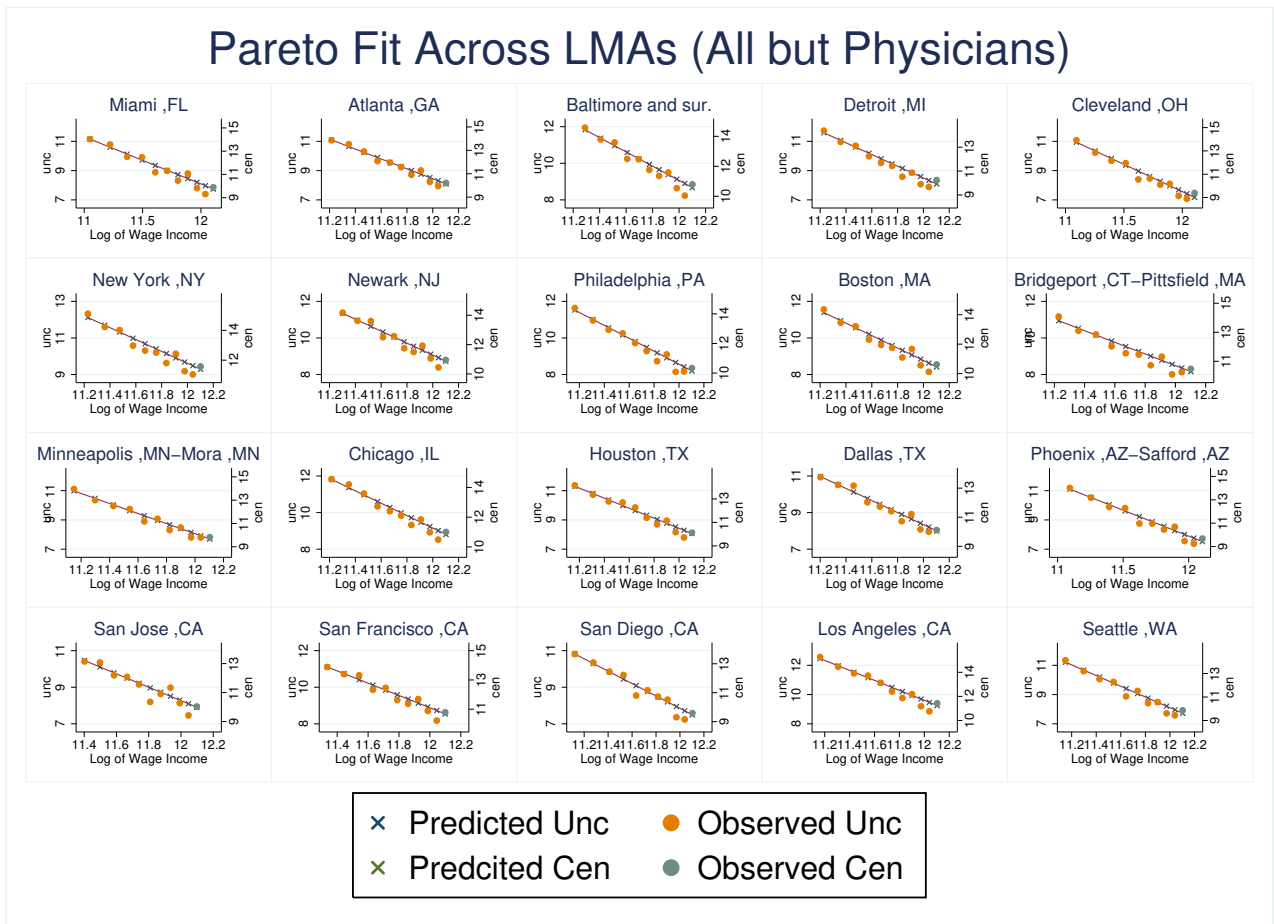
The paper uses the assumption of Pareto for physicians on the LMA-year-occupation level, for LMA-year for the general population and for occupation-year level for the top 20 occupations. Figure 6 in the main text shows the fit with Pareto distribution for the biggest LMA for the whole distribution and for physicians specifically. Figures D.1 and D.2 show analogous figures for the 20 biggest labor market areas for physicians and for the all other occupations than physicians both for the year 2000. Whereas the general population fits the Pareto assumption remarkably well, there is more noise around the line for the physicians, though no systematic deviation.

Figure D.1: Fit to the Pareto Distribution for general income distribution for Physicians for 20 biggest labor market areas for 2000 (using top 65 per cent of uncensored observations)



Notes: Using top 65 per cent of uncensored observations

Figure D.2: Fit to the Pareto Distribution for general income distribution excluding Physicians for 20 biggest labor market areas for 2000 (using top 10 per cent of uncensored observations)



Notes: Using top 10 per cent of uncensored observations

Table D.1: Top occupations and income inequality ($1/\alpha$)

Occupation	1980 (pred. 95/90)	$1/\alpha$		2014 (pred. 95/90)
		1990	2000	
Chief executives and public administrators	0.24 (1.18)	0.34	0.65	0.57 (1.48)
Financial managers	0.32 (1.25)	0.43	0.48	0.52 (1.44)
Managers and specialists in marketing, advertising, and public relations	0.30 (1.23)	0.33	0.36	0.37 (1.29)
Managers in education and related fields	0.19 (1.14)	0.24	0.23	0.29 (1.22)
Managers and administrators, n.e.c.	0.43 (1.34)	0.45	0.36	0.38 (1.30)
Accountants and auditors	0.27 (1.21)	0.32	0.38	0.44 (1.35)
Not-elsewhere-classified engineers	0.22 (1.16)	0.23	0.24	0.23 (1.17)
Computer systems analysts and computer scientists	0.16 (1.12)	0.21	0.25	0.25 (1.19)
Physicians	0.47 (1.39)	0.78	0.55	0.62 (1.54)
Registered nurses	0.17 (1.13)	0.17	0.20	0.23 (1.17)
Subject instructors (HS/college)	0.20 (1.14)	0.24	0.28	0.33 (1.26)
Lawyers	0.42 (1.34)	0.53	0.58	0.58 (1.49)
Computer software developers	0.18 (1.13)	0.19	0.23	0.24 (1.18)
Supervisors and proprietors of sales jobs	0.40 (1.32)	0.45	0.44	0.44 (1.36)
Insurance sales occupations	0.42 (1.34)	0.50	0.52	0.58 (1.49)
Salespersons, n.e.c.	0.35 (1.27)	0.40	0.39	0.42 (1.34)
Supervisors of construction work	0.30 (1.23)	0.33	0.29	0.30 (1.23)
Production supervisors or foremen	0.20 (1.15)	0.20	0.26	0.29 (1.23)
Truck, delivery, and tractor drivers	0.20 (1.15)	0.22	0.24	0.26 (1.20)
Military	0.28 (1.21)	0.25	0.28	0.25 (1.19)

Notes: Estimates of $1/\alpha$ for top 20 occupations using top 10 per cent of population

Table D.2: Top occupations in 1%, 5%, 10%, for year 1980 and 2014

rank	1980			2014		
	top 10 pct	top 5 pct	top 1 pct	top 10 pct	top 5 pct	top 1 pct
1	Managers and administrators, n.e.c.	Managers and administrators, n.e.c.	Managers and administrators, n.e.c.	Managers and administrators, n.e.c.	Managers and administrators, n.e.c.	Physicians
2	Salespersons, n.e.c.	Salespersons, n.e.c.	Physicians	Chief executives and public administrators	Physicians	Managers and administrators, n.e.c.
3	Production supervisors or foremen	Production supervisors or foremen	Salespersons, n.e.c.	Computer software developers	Chief executives and public administrators	Chief executives and public administrators
4	Truck, delivery, and tractor drivers	Managers and specialists in marketing, advertising, and public relations	Managers and specialists in marketing, advertising, and public relations	Physicians	Lawyers	Lawyers
5	Managers and specialists in marketing, advertising, and public relations	Physicians	Lawyers	Lawyers	Salespersons, n.e.c.	Supervisors and proprietors of sales jobs
6	Supervisors of construction work	Supervisors and proprietors of sales jobs	Supervisors and proprietors of sales jobs	Salespersons, n.e.c.	Supervisors and proprietors of sales jobs	Salespersons, n.e.c.
7	Supervisors and proprietors of sales jobs	Truck, delivery, and tractor drivers	Insurance sales occupations	Supervisors and proprietors of sales jobs	Computer software developers	Financial managers
8	Physicians	Lawyers	Production supervisors or foremen	Computer systems analysts and computer scientists	Managers and specialists in marketing, advertising, and public relations	Other financial specialists
9	Accountants and auditors	Supervisors of construction work	Real estate sales occupations	Registered nurses	Financial managers	Managers and specialists in marketing, advertising, and public relations
10	Electrical engineer	Financial managers	Airplane pilots and navigators	Accountants and auditors	Accountants and auditors	Accountants and auditors

Notes:

E Additional results

E.1 Additional Regressions for other occupations

We perform an analysis like that of Tables 7 and 8 for nurses, College professors and Real Estate agents (occupation code 254). Real Estate agents are censored at around top 7 per cent and we use top 20 per cent uncensored observations.

Finally, we show that income inequality for chief executives and public administrators positively predict the income inequality for secretaries in Table E.3.

E.2 Robustness Checks for Physicians

We perform robustness checks for the the regression in Table 7. In particular, Table E.1 shows the regression for different cut-offs. The parameter estimate is generally not far from 1 and remains significant at the 10% level throughout the regressions. Table E.2 shows that the choice of how many LMAs to include does not affect the parameter estimate much.

Table E.1: IV Regressions for Physicians for different cut-offs of Pareto Distribution

cut-off	(1) 35	(2) 40	(3) 45	(4) 55	(5) 65	(6) 75
log(1/ $\alpha(-o)$)	1.30*** [0.25, 2.42]	1.48** [0.06, 2.23]	1.41* [-0.35, 2.86]	1.07 [-0.44, 2.00]	1.17* [-0.12, 2.41]	1.82 [-0.34, 3.67]
logpop	0.09 [-0.09, 0.22]	0.13 [-0.11, 0.26]	0.17* [-0.02, 0.34]	0.17* [-0.02, 0.35]	0.20** [0.02, 0.47]	0.31** [0.02, 0.48]
log(income)	0.06 [-0.11, 0.17]	-0.03 [-0.23, 0.12]	-0.25*** [-0.45,-0.05]	-0.31*** [-0.56,-0.17]	-0.32*** [-0.66,-0.16]	-0.25** [-0.71,-0.06]
1990	-0.07 [-0.19, 0.11]	-0.04 [-0.14, 0.15]	-0.20 [-0.70, 0.03]	-0.38*** [-0.92,-0.21]	-0.41*** [-0.97,-0.23]	0.20** [0.08, 0.60]
2000	-0.20 [-0.40, 0.12]	-0.17 [-0.33, 0.19]	-0.08 [-0.36, 0.06]	-0.17** [-0.45,-0.08]	-0.16** [-0.45,-0.04]	0.00 [-0.17, 0.66]
2014	-0.14 [-0.37, 0.22]	-0.09 [-0.27, 0.31]	-0.13*** [-0.19,-0.10]	-0.21*** [-0.28,-0.18]	-0.25*** [-0.34,-0.21]	0.18* [-0.03, 0.94]
Observations	1,012	1,012	1,012	1,012	1,011	1,011

Bootstrapped standard errors based on 100 draws, stratified at the occupation/year/labor market level. 95 pct confidence interval in square parentheses. Income is average wage income for those with positive income o refers to occupation of interest * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$

Table E.2: IV Regressions for Physicians for different number of LMAs

LMAs	(1) 100	(2) 150	(3) 200	(4) 253	(5) 300	(6) All
log(1/α(-o))	1.70*** [0.76, 2.45]	1.35** [0.27, 2.01]	0.83* [-0.16, 1.57]	1.30*** [0.25, 2.42]	1.21** [0.09, 2.10]	1.11** [0.05, 1.87]
logpop	0.17 [-0.03, 0.33]	0.07 [-0.12, 0.20]	0.02 [-0.14, 0.17]	0.09 [-0.09, 0.22]	0.12 [-0.05, 0.21]	0.09 [-0.05, 0.21]
log(income)	-0.04 [-0.22, 0.15]	0.03 [-0.16, 0.17]	0.09 [-0.08, 0.19]	0.06 [-0.11, 0.17]	0.07 [-0.09, 0.22]	0.08 [-0.11, 0.24]
1990	0.16 [-0.22, 0.39]	0.13 [-0.25, 0.26]	0.05 [-0.30, 0.24]	-0.07 [-0.19, 0.11]	0.14 [-0.26, 0.32]	0.11 [-0.24, 0.26]
2000	0.09 [-0.11, 0.22]	0.07 [-0.13, 0.14]	0.02 [-0.16, 0.11]	-0.20 [-0.40, 0.12]	0.07 [-0.14, 0.18]	0.05 [-0.12, 0.15]
2014	-0.02 [-0.10, 0.03]	-0.04** [-0.10,-0.01]	-0.05*** [-0.11,-0.02]	-0.14 [-0.37, 0.22]	-0.06*** [-0.13,-0.03]	-0.08*** [-0.13,-0.04]
Observations	400	600	800	1,012	1,200	1,573

Bootstrapped standard errors based on 100 draws, stratified at the occupation/year/labor market level. 95 pct confidence interval in square parentheses. Income is average wage income for those with positive income o refers to occupation of interest * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$

Table E.3: OLS regressions for secretaries on Chief executives and public administrators for 2000 and 2014

	(1)	(2)	(3)	(4)
	Secretaries	Secretaries	Secretaries	Secretaries
Chief executives and public administrators	0.178*** (4.24)	0.177*** (4.28)	0.136** (2.02)	0.136** (2.00)
2014		-0.0953*** (-4.50)	-0.0968*** (-4.64)	-0.0892 (-1.21)
Log of Inc.				-0.0446 (-0.18)
Log of Pop.				0.0391 (0.19)
Observations	769	769	769	769
Adjusted R^2	0.022	0.046	0.095	0.090

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Regressions limited to 2000 and 2014 due to insufficient information on CEOs in 1980 and 1990. Weighted by number of secretaries by LMA. For 8 observations or more. Column (I) is univariate OLS, Column (II) includes time dummy for 2014, Column (III) further includes labor market area fixed effects and Column (IV) controls for average wage income as well as population.

F Construction of Data on Labor Market Areas

The publicly available data from IPUMS gives information on “country group” in 1980 and “Public Use Microdata Area” (PUMA) for 1990 and onward. We wish to assign these to labor market areas. Dorn (2009) uses a probabilistic approach using the aggregate correspondence between county groups/PUMAs and counties and counties and commuting zones and creates a “crosswalk” assigning weights for each country group in 1980 to 1990 commuting zones and for each PUMA to 1990 commuting zones. If a given county group or PUMA is assigned to multiple commuting zones we “split” all individuals in the county group or PUMA and give each weights from the crosswalk. The IPUMS data from 2012 onward uses the PUMA2010 (updated from the 2010 federal census) and we construct a new crosswalk along the same lines as Dorn (2009). Counties are very stable across town and we manually correct for county changes between 2000 and 2010. Finally, since our unit of analysis is labor market areas we use Missouri Census Data Center (<http://mcdc.missouri.edu/websas/geocorr2k.html>) to aggregate commuting zones into labor market areas. Each commuting zone is uniquely assigned to a labor market area. If a single individual had been split into two commuting zones within the same labor market area using Dorn’s algorithm we combine the two into one observation aggregating their weights. Figure F.1 shows the labor market areas for 1990.

Labor market areas for the United States, 1990



Figure F.1: This map shows the 1990 Labor Market Areas, which we use as the local market geographic unit throughout the paper.