Technical change and superstar effects: evidence from the roll-out of television

Monograph:

König, Felix (2019) Technical change and superstar effects: evidence from the roll-out of television. CEP Discussion Papers. Centre for Economic Performance, LSE, London, UK.

ISSN 2042-2695

CENTRE *for* ECONOMIC PERFORMANCE

CEP Discussion Paper No 1663

November 2019

Technical Change and Superstar Effects: Evidence from the Roll-Out of Television

Felix Koenig

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

UKRI Economic and Social Research Council

Abstract
"Superstar effects" generate large compensation differentials among similarly talented individuals. Are superstar effects amplified by technological innovations that extend the scale over which talent is deployed? I test this idea in the market for entertainers, using the roll-out of television as a natural experiment which provides clean variation in a scale-related technological change. The launch of a local TV station increases top entertainers' incomes, resulting in a twofold increase in top-percentile income share, while reducing employment and incomes of lower-level talents. These results show clear evidence of superstar effects and are inconsistent with canonical models of skill-biased technological change.

Key words: Superstar Effect, inequality, top incomes, technical change
JEL Codes: J31; J23; O33; D31

This paper was produced as part of the Centre's Labour Markets Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

I thank Daron Acemoglu, Josh Angrist, David Autor, Jan Bakker, Oriana Bandiera, Jesse Bruhn, Ellora Derenoncourt, Jeremiah Dittmar, Thomas Drechsel, Fabian Eckert, Horst Entorf, George Fenton, Andy Ferrara, Torsten Figueiredo Walter, Xavier Gabaix, Killian Huber, Simon Jaeger, Philipp Kircher, Camille Landais, Matt Lowe, Alan Manning, Ben Moll, Niclas Moneke, Markus Nagler, Barbara Petrongolo, Steve Pischke, David Price, Arthur Seibold, Marco Tabellini, Lowell Taylor, Catherine Thomas, Joonas Tuhkuri, Anna Valero and John van Reenen.
Felix Koenig, Princeton University, IZA and Centre for Economic Performance, London School of Economics.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

1 Introduction

Income inequality has risen sharply in recent decades, particularly at the top of the distribution. In the US, for instance, the income share of the top 1% of earners increased from about 8% in 1970 to over 20% today. What explains such large income differences and why are they growing, particularly at the top of the distribution? Many economists link such trends to technical change. An influential literature has highlighted the role of "skill-biased technical changes" (SBTC) that increase in the demand for high skilled workers, relative to low skilled workers and potentially explains part of the 20th-century shifts in the US skill premium.[1] An alternative explanation for rising inequality at the top end focuses on increasing "superstar effects," which generate large wage differences among workers with nearly identical talent, particularly at the top of the distribution.[2] Crucial to the magnitude of such effects are technologies that expand the *scale* over which talent is deployed. In recent years, progress in communication technologies (and other similar technologies) has increased the potential scale of many productive activities, which plausibly amplified superstar effects and thereby generated growth in income at the very top. Recent studies show a strong correlation of top income growth and rising production scale in many sectors of the economy.[3] Beyond such suggestive evidence, there is, however, very little empirical work that directly tests whether scale related technical changes are generating superstar effects.

I use a natural experiment in the entertainment sector to test whether scale-related technologies amplify superstar effects, and find strong evidence for such effects. Specifically, I study the roll-out of television stations in the US that sharply expanded the scale of entertainment shows. Television filming initially started city by city and a staggered deployment of TV stations leads to variation in exposure across local labor markets. I show that the launch of a local TV station increases top entertainers' incomes, resulting in a twofold increase in top-percentile income share, while reducing employment and incomes of lower-level talents.

Technological change provides a potentially powerful explanation for recent changes in inequality, but a lack of exogenous variation in technical change makes

[1]See, e.g., Bound and Johnson (1992), Katz and Murphy (1992) and the many papers that followed.

[2]The superstar theory is developed in Rosen (1981), Tinbergen (1956)

[3]See work on superstar effects by Gabaix and Landier (2008), Terviö (2008) for CEOs, Garicano and Hubbard (2009) for lawyers, Célérier and Vallée (2019), Kaplan and Rauh (2010) finance professionals, and Krueger (2005) for entertainers.

it difficult to test such effects. Card and DiNardo (2002), Lemieux (2006) famously stress that recent inequality growth coincided simultaneously with accelerated technical change and also with reforms to labor market institutions, thereby making it difficult to isolate the effect of technical change. In the context of SBTC, this identification challenge has been addressed by new research focusing on the impacts of specific technologies (Akerman, Gaarder, & Mogstad, 2013; Bartel, Ichniowski, & Shaw, 2007; Michaels & Graetz, 2018). The objective of my paper is to provide comparable well-identified empirical evidence relevant to superstar effects.

The first key contribution of the paper is to develop a tractable model of superstar effects, and derive testable predictions that distinguish such effects from canonical models of technical change (including SBTC). A well-known cross-sectional prediction of superstar effects is that small talent differences become amplified into large income differences, but a test of this prediction requires a credible cardinal measure of talent. I overcome this challenge by looking at dynamic predictions of a superstar model – focusing on changes in inequality predicted by the model. I show that superstar effects are amplified by a specific type of technical change, scale-related technical change (SRTC), which increases market reach. These technologies enable the most talented workers in the profession, the "superstars," to attract a greater share of customers, thereby increasing income concentration at the top. At the same time, the model predicts falling returns and employment loss at lower talent levels. This superstar pattern is different from a large class of alternative mechanisms, including canonical SBTC models, where technical change leads to rising labor demand and rising returns across the talent distribution, proportional income gains and rising employment.

The second contribution of the paper is to implement a test of the superstar model in the entertainment sector. The staggered roll-out of television, an obvious SRTC, provides a near-ideal testing ground for superstar effects within this sector. Television of course eventually led to a vast increase in the scalability of entertainers' performances, but this transformation took place in stages. Before TV, entertainers' performances could typically be watched by only a few hundred individuals in local venues. When TV was first introduced in the US in the 1940s, technical constraints required television filming to take place near broadcast antennas, and television filming thus occurred in multiple local labor markets. This practice meant that there were many local TV stars scattered across the US. By the mid-1950s, the introduction of videotaping led to national scalability of entertainment. Today superstars have a worldwide reach and are among the highest-paid individuals in the economy.[4]

By using a natural experiment, the city-by-city introduction of TV, I sidestep two problems faced by the literature on technical change and inequality. A first challenge is that the local availability of technologies is rarely observed directly. Instead, researchers have to make do with proxies for the technical change (e.g., by using variation at the industry or occupational level), which makes it difficult to identify the effect of technical change in the face of spurious trends, for instance arising from deregulation or pay-setting norms.[5] In my setting, by contrast, variation arises within a given industry across distinct labor markets. To make the TV roll-out traceable, I geocode locations of TV stations from archival records and then use this variation in a differences-in-differences analysis among local entertainer labor markets. This approach allows me to control for confounding industry-level time trends and thus distinguish the effects of technology from confounding aggregate or industry-specific trends in deregulation or pay-setting norms. A second challenge for this literature is the endogenous adoption of technologies.[6] Technology adaption that responds to local labor market conditions creates a simultaneity problem which complicates efforts to identify causal impacts of the technical innovation. I avoid this problem by exploiting government deployment rules that generate variation in TV adoption, independent of local demand shocks.[7]

I find strong evidence for rising superstar effects. With the launch of a TV station, incomes of top entertainers grow sharply, especially at the very top of the distribution. As a result, the share of income going to the top 1% nearly doubles. Such gains arise only at the top of the distribution; mid-level talents, by contrast, lose out. The results show a decline in the number of mid-paid entertainment jobs and an overall contraction in employment in the industry by approximately 13%. These demand shifts are also reflected in expenditure records that show a substantial decrease in

[4]Within the top 0.1% highest-paid Americans, only finance professionals and entrepreneurs receive higher incomes than entertainers. Despite being a relatively small sector, entertainers contribute more to top income shares than medical professionals or CEOs of publicly traded companies and have a similar contribution to engineers. Based on Table 3a & Table 7a of Bakija, Cole, and Heim, 2012 and ExecuComp records on the compensation of CEOs of publicly traded companies.

[5]Famous proxies for technical change based on occupation or industry variation is presented in, e.g. Acemoglu and Restrepo (2019), Autor and Dorn (2013), Goos, Manning, and Salomons (2014)

[6]For endogenous technical change see for instance Acemoglu (1998).

[7]The government selected locations based on predetermined local characteristics which make the deployment of TV unresponsive to local demand shocks. In addition, I can evaluate the validity of this approach by exploiting an unexpected interruption of the TV roll-out, during which several locations narrowly miss out on planned TV launches. In such nearly-treated places, we see no effects, adding confidence to the assumption that the roll-out is unrelated to spurious local demand shocks.

spending at ordinary entertainment events, but increases in audiences and returns for the most successful shows. In short, scale related technical change has the effect predicted by superstar theory and moved the industry toward a winner-takes-all extreme.

Recent research stresses that superstar effects could play an important role in shaping the modern economy. Such studies suggest that superstar effects could rationalize the observed shifts in the distribution of earnings, especially as a driver of rising top income inequality (see, e.g., Brynjolfsson & McAfee, 2011; Guellec & Paunov, 2017; Kaplan & Rauh, 2013), and may also influence the regional location of economic activity (Eckert, Ganapati, & Walsh, 2019). These suggestive results highlight the need for a clearer understanding of SRTC and superstar effects in the modern economy. Understanding the drivers of income inequality is also highly relevant for policy-making. Studies on the level and progression of taxes, for instance, highlight that the appropriate policies depend crucially on understanding the forces that drive income inequality (Scheuer & Werning, 2017).

2 The Superstar Model

In this section I present a tractable framework to study superstar effects and derive testable predictions that distinguish such effects from a wide range of alternative mechanisms. A first key ingredient of superstar effects are workers with different and unique levels of talent who are matched with heterogenous tasks. In the context of entertainment we can think of workers as actors and of tasks as shows. The position in the distribution is given by the inverse CDF, denoted by $p^t \equiv P(t > t_p) = h(t_p)$ and $p^s \equiv P(s > s_p) = g(s_p)$, for talent t and show audience size s respectively and I assume that these distributions are continuous.[8] The most talented entertainer is ranked top of the talent distribution and thus has a value of $p^t = 0$, this rank value increases towards one . In the production process actors are matched with a stage and the revenue of a matched pair is given by $Y(s,t)$ with standard properties: $Y_s > 0, Y_t > 0, Y_{tt} < 0$. A second crucial assumption for superstar effects is comparative advantage of more talented actors in bigger shows ($Y_{st} > 0$) (see Sattinger, 1975). This assumption

[8] Continuity is not essential to the model; with jumps in the distribution one actor (or stage) would be discretely better than the next, which would generate monopoly power and lead to match specific rents. The exposition here abstracts from this and I return to the topic below.

leads to positive assortative matching of actors and shows in equilibrium.[9]

The equilibrium of this model is characterized by three conditions (derived in Appendix 9.1). These conditions respectively guarantee positive assortative matching, incentive compatibility and market clearing. Together these conditions pin down the equilibrium values of wages $w(t^*)$, equilibrium assignment $\sigma(t^*)$, and output prices π^*:

$$p^t = p^s(\sigma(t^*)) \tag{1}$$

$$w'(t^*) = Y_t(\sigma(t^*), t^*) \tag{2}$$

$$\int_0^1 h'(t)Y(\sigma(t), t)dt = D(\pi^*) \tag{3}$$

Equation 1 is positive assortative matching (PAM) between actors and shows and implies that the best actor works on the biggest stage. More generally, the two match partners are at the same percentiles of their respective size and talent distributions. Equation 2 is the incentive compatibility (IC) constraint and ensures that there are no profitable deviations from the assignment in 1. Note that equation 2 holds that equilibrium wages increases in line with the marginal product of workers. The equilibrium is therefore perfectly competitive in the sense that there are no match specific rents. There is also an implicit participation constraint that ensures that the lowest wages at percentile $\bar{p}$ is above the outside options (w^{res}): $w(\bar{p}) \geq w^{res}$. For now assume that this is always satisfied and labor supply is thus inelastic. The final condition, equation 3, ensures that demand for talent $D(\pi)$ equals the total of units of talent supplied at the equilibrium prices π^*.[10]

The resulting superstar wage distribution has been the key focus in the literature and distributional results have been derived for the general case. Here, I will focus on a closed form solution to illustrate the key mechanics of the model. Assume therefore that talent t and show size s follow Pareto distributions with shape parameters α and β for actor talent and show size respectively ($p^t = t_p^{-\frac{1}{\alpha}}$ and $p^s = s_p^{-\frac{1}{\beta}}$). A larger value of the shape parameter implies greater dispersion in the talent and

[9]Alternatively, papers have a related log-supermodularity assumption and arrive at the same results. Neither assumption is stronger in the sense that one implies the other.

[10]We assume that workers supply one unit of labor inelastically, an assumption that is relaxed below.

size respectively. Moreover, assume that the production function is Cobb-Douglas $F(s,t) = \pi \left[s^{(1-\gamma)} t^{\gamma}\right]^{\phi}$, where ϕ determines the scalability of production ($\phi > 0$) and $\gamma \, \epsilon \, (0,1)$. These assumptions lead to the wage distribution of the superstar economy:[11]

$$p^w = \lambda {w_p}^{-\frac{\xi}{\alpha}} \tag{4}$$

as before p^w indicates an inverse percentile, here in the wage distribution. Wages follow a Pareto distribution, with the shape parameter $\frac{\alpha}{\xi}$, where $\xi \equiv \frac{1}{\phi} \frac{\alpha}{\gamma\alpha + (1-\gamma)\beta}$ and scale parameter $\lambda \equiv (\gamma\phi\pi)^{\xi/\alpha}$. Notice that the wage distribution looks similar to the distribution of talent. The shape parameter of the talent distribution is α, hence wages are more dispersed than talent if $\xi < 1$. For small values of ξ the superstar model therefore produces large wage differences, even if talent differences are small. This "talent amplification effect" has been the focus of much early literature on superstar models (e.g. Rosen, 1981; Sattinger, 1975; Tinbergen, 1956).[12] It is however difficult to test this effect because a test would require a cardinal measure of talent. The lack of objective talent units makes this prediction indistinguishable from an alternative model without the talent amplification effect and extremely rare talent.

2.1 The Effect of Technical Change

We can however distinguish superstar effects from a wide range of other labor demand models by looking at the effect of technical change. The distinction becomes visible during periods of scale related technical change (SRTC), a technical change that makes production more scalable ($\tilde{\phi} = \kappa\phi$ with $\kappa > 1$). We can see the effect by looking at the change in the share of jobs with wage w, denoted by $g_e(w)$. This change is derived by differentiating equation 4 and comparing the resulting distribution before and after the shift in ϕ:

$$g_e(w) = \frac{\tilde{\lambda}\tilde{\xi}}{\lambda\xi} w^{\frac{(\kappa-1)}{\kappa}\frac{\xi}{\alpha}} - 1 \tag{5}$$

[11]Condition 1 becomes $\sigma(\hat{t}) = \hat{t}^{\frac{\beta}{\alpha}}$ and 2: $w'(\hat{t}) = \gamma\phi\hat{\pi}\hat{t}^{(\frac{1}{\xi}-1)}$. Integrating this wage schedule and normalizing the constant of integration, or workers' outside option, to zero we arrive at the wage distribution in the superstar economy.

[12]In what follows I focuses on cases where this amplification effect holds. This is the case as long as large show venues are scarce enough to overcome diminishing returns to scale as we move up in the distribution (aka if $\frac{\beta}{\alpha} > \frac{1-\gamma}{\gamma}$).

where "$\tilde{x}$" indicating new values. Notice that $g_e(w)$ is biggest for high values of w, SRTC thus leads to a growing fraction of top paid actors. This captures the best known implication of superstar effects: sharp gains at the top of the distribution. Moreover, such effects are biggest at high values of w and are diminishing as we move away from the top of the distribution. A second effect operates through the drop in λ, the shape parameter of the wage distribution. The greater availability of stars, puts downward pressure on all wages in the distribution by reducing π and consequently λ and leads to declining share of mid income jobs.[13] For non star jobs this effect outweigh the gains from greater scalability and turns previous mid income jobs into low paid jobs, in the limit such losses lead to a single superstar that serves the entire market. The impact of Superstar Effects across the wage distribution are summarized in Figure 1. High paid jobs emerge at the very top, while growth rates diminish as we move away from the stars of the profession. Towards the middle of the distribution mid-income jobs are disappearing as mediocrely talented workers pay is declining, leading to growing share of workers with low pay. The impact of superstar effects across the wage distribution is therefore U-shaped. To map this model to the empirical setting, we additionally want to allow workers to exit the sector, reflecting that the empirical setting studies a partial equilibrium set-up. The wage in the outside sector is given by w^{res} and workers with wages below this level leave the sector. This leads to a participation threshold, denote the level of the lowest participating talent $\bar{p}$.

The distributional consequences of Superstar Effects can be summarized by four testable propositions (derived in Appendix 9.3):

Proposition 2.1. *Superstar Effects lead to*

a) Top wage growth: For two percentiles at the top of the wage distribution $p' > p$ the growth rate g_e increases as we move up in the distribution: $g_e^{p'} > g_e^p$

b) Fractal inequality: For top income shares (s_p) at two percentiles p pay differences increase: $\tilde{s}_{1\%}/\tilde{s}_{10\%} > s_{1\%}/s_{10\%}$

c) Adverse effects for lesser talents: Employment at mid paid levels declines as $g_e < 0$ when $w \to 1$

[13]Notice that if π is unchanged (ie if demand for entertainment is perfectly inelastic), λ would rise. I assume that demand is sufficiently elastic ($1 - \varepsilon < (\gamma\phi\pi)^{\kappa-1}$) to rule this case out, for an alternative approach to generate this result see Rosen, 1981. Even without these cannibalizing effects, the previous results on relative gains at the top hold. This is for instance the case in Gabaix and Landier, 2008.

d) Employment loss: For a given outside option w^{res} and corresponding participation threshold $\bar{p}$ superstar effects imply $\tilde{\bar{p}} < \bar{p}$

The first result in *a*) is disproportionate gains at the top of the income distribution. The second result in *b*) focuses on growing income dispersion *within* the top income tail and captures that moving up a rank in the talent distribution becomes more valuable. As a result, a growing proportion of the the income earned by the top 10% is earned by the top 1% and consequently the ratio of the two income shares ($s_{1\%}/s_{10\%} \uparrow$) increases, also known as growing "fractal inequality". The third result in *c*) highlights that gains at the top come at the expense of other entertainers. Technical progress allows stars to steal business of lesser talents, leading to falling demand for such lower talent entertainers. The final results in *d*) captures the winner takes all nature of superstar effects. Employment falls when the stars' growing market reach reduces returns of other workers below the reservation utility and thus pushes them out of the market.

Such superstar effects are different from canonical models of labor demand. Canonical labor demand models, including SBTC, feature skill groups of perfectly substitutable workers. Among perfectly substitutable workers the law of one price limits growth in wage dispersion and makes wage growth proportional to talent, at odds with results *a*) and *b*) above.[14] A second distinctive feature of Superstar Effects are the losses incurred by large parts of the distribution. Caselli and Manning (2019) show that such losses are at odds with canonical labor demand models, where technical change shifts labor demand outward and increases both wages and employment, at odds with the Superstars results *c*) and *d*) above.[15]

In summary, we can distinguish the superstar effects by taking 4 predictions to the data: a) disproportional wage growth at the top, b) growing pay inequality among top earners c) decreasing wages for mediocre workers, and d) falling employment.

3 Data and Setting

The entertainment sector in the mid 20th century provides sharp variation in production scalability and a rare opportunity to simultaneously observe the locations

[14] Appendix 9.4 shows these limitations of SBTC and illustrates the extensions that are required to make the model match superstar effects.

[15] Models where technical change improves the productivity of capital instead of workers (e.g. routinization) may lead to wage and employment losses if capital is a substitute for workers.

of technical change, the resulting shift in market reach and labor market outcomes. I combine records from archival sources to measure all three and isolate plausibly exogenous variation to test for superstar effects.

3.1 Production Technology

With the launch of TV in the 1940s and 1950s entertainment shows began to reach mass audiences. Broadcasting antennas relayed shows via airwaves to households in the transmission radius. At the time TV sets were not as widely available as today, nevertheless the audiences of TV shows was substantially bigger then at local live shows and audiences available to entertainers multiplied. Pioneering TV filming, unlike today, was predominantly live and recorded locally near the broadcasting station. Multiple local stations were therefore filming TV shows simultaneously across the country. The fragmentation of filming was the result of technological and regulator constraints of the early TV period. Most importantly, the lack of infrastructure to transmit shows from station to station inhibited relaying the same shows across the US (see Sterne (1999) for a detailed account). Moreover, recording technologies were in their infancy and resulted in poor image quality, making them a poor substitute for local live material.[16] Finally, regulation restricted studio locations and specified that "the main studio be located in the principal community served" (FCC annual report 195) and TV filming thus occurred in places scattered across the country.

To trace the location of filming I geocode the location of studios and match them to local labor markets. Previous studies on television have focused on the other side of the market, the consumers of TV shows and traced TV signal transmission to test the effects of watching TV on political and educational outcomes (e.g., Gentzkow, 2006; Gentzkow and Shapiro, 2008). This study instead digitizes new administrative records on the location of TV filming to study changes in labor demand for entertainers. Moreover, I sharpen the identification strategy with novel data on the licensing process. The data come from archival records of the "Annual Television Factbooks," and I compute the number of stations filming in each US local labor market.[17]

[16]Non-local content had to be put on film and shipped to other stations, where a mini film screening was broadcast live, this was known as "kinescope."

[17]I assume that all stations were filming locally at that time. A handful of stations are an exception and operated a local network. This was rarely feasible because the technical infrastructure was still in its infancy. In my main specifications I code all members of such networks as treated to avoid potential endogenous selection of filming locations within the network.

Figure 2 shows where TV filming took place in 1949, a year with Census wage data. Exposure to TV filming varied substantially across local labor markets, and I exploit this variation in a differences-in-differences analysis.

The key identification assumption is that TV launches are unrelated to local demand shocks. Central to this assumption are the deployment rules of TV stations. For years when such rules are available archival records lend credibility to this assumption.[18] Administrators drew up lists that ranked locations in terms of priority and worked through the ranking. Priority is based on pre-determined location characteristics and thus based on factors that are unresponsive to local demand shocks. This guards against the most glaring endogeneity concerns from endogenous adaption. We may however worry that there are more complex spurious links between the assignment and labor market changes.

An unplanned halt in licensing leads to additional quasi-experimental variation in TV that allows me to probe if the roll-out is indeed unrelated to local demand changes. The hold-up gives rise to a group of locations that would have received TV but narrowly miss out due to the regulator shut-down. I use such blocked locations to test if the roll-out is related to spurious demand shocks. The principal reason for the interruption in the roll-out was an error in the FCC's airwave propagation model. This model was used to delineate interference-free signal catchment areas, but the error implied that signal interference occurred among neighboring stations. To avoid a worsening of the situation, the FCC put all licensing on hold and ordered a review of the model. Previous studies noted this hold up, but lacked the regulator records to distinguish locations that were held up by the FCC from late adapters. I collected new FCC records to distinguish the two groups and Figure 2 shows the affected labor markets. Licensing only resumed in 1952 and delayed the onset of TV by at least four years in the held up locations.[19]

A further attraction of this setting is that we observe the end of local TV filming. Local shows are eventually superseded by centralized productions. The Ampex videotape made shows from outside the local labor market a close substitute for local live shows and led to the concentration of TV production in two hubs, Los Angeles and New York and the demise in other locations.[20] The videotape was presented

[18]The target areas were at least 50 miles from the nearest station and in 1952 ordered by 1950 population.

[19]Initially the freeze was expected to last about a year. However, the review was delayed to ensure compatibility with arising new transmission technologies (UHF and color transmission).

[20]This trend was also helped by the contemporaneous roll-out of coaxial cables that allowed

in 1956 and immediately over 70 videotape recorders were ordered by TV stations across the country. The same year, CBS started to use the technology, and the other networks followed suit the next year and thus led to the rapid decline of local filming. This removal of local filming gives rise to another identification check and we can test if the local effects disappear. For this test I control for places where filming centralizes and to avoid an endogenous control problem, I use a pre-determined measure of production costs as control. This variable picks up location incentives that come from permanent regional characteristics and is based on the share of 1920 movies produced in each local labor market, gathered from records of the "Internet and Movie Database" (ImDB)

3.2 Labor Market Data

Labor market data come from samples of the of the US Census micro-data files (1930-1970). I focus on five entertainment occupations that benefited from the introduction of TV: actors, athletes, dancers, musicians and entertainers not elsewhere classified. For each of them I compute labor market outcomes for the 722 local labor markets that span the mainland US.[21] The Census first collected wage data in 1940 and in all years asked about the previous year, wages reported in 1940 thus refer to 1939. Data for the full distribution of wages is reported in 1940, but from 1950 onward top coding applies. Fortunately, the top code bites above the 99th percentile of the wage distribution and up to that threshold, detailed analysis of top incomes is possible. A first set of variables looks at entertainers' position in the US wage distribution. This follows Chetty, Hendren, Kline, and Saez (2014) in measuring inequality by ranking entertainers' wages relative to the overall labor force and thus focuses on a metric that is scale independent, making comparisons over time easier. This measure also side-steps top-coding issues, as the share of workers with a wage above a threshold, say the 99th percentile, can be computed. The share of top paid entertainers, say those whose wage falls in the top 1% of the US wage distribution ($D^{US1\%} = 1$), is thus given by:

$$p^{99}_{m,t} = \frac{\sum_i E_{i,m,t} \cdot D^{US1\%}}{\overline{E_t}}$$

which is simply the ratio of top-earning entertainers in market m at time t and the number of entertainers in a labor market. One concern with this outcome is that

producers to relay live shows from station to station.

[21] I follow Autor and Dorn (2013) and define local labor markets based on commuting zones.

fluctuations in the denominator lead to spurious effects. To address this challenge, I hold the denominator fixed at the average national level, which turns the division into a normalization.[22] An alternative solution studies outcomes at a per-capita level, which is also unaffected by swings in entertainer employment. I additionally compute such per capita measures, that use local population as denominator. A further outcome of interest is the share of income going to top percentiles in entertainment. To compute such shares I have to take a stance on the pay distribution beyond the top code. To do this I follow the literature in using Pareto approximations.[23] Moreover, I compute additional outcome measures, including log employment and geographic mobility of entertainers (see Appendix 10.3 for details).[24]

3.3 Market Size

The entertainment setting offers a unique opportunity to quantify what market workers serve. Such data comes from archival records on audiences and revenues of live and TV shows. For live shows I use the 1921 "Julius Cahn-Gus Hill theatrical guide," which aims to provide "complete coverage of performance venues in US cities, towns and villages."[25] For TV shows I compute the number of TV households in a station's signal catchment area using signal data from Fenton and Koenig (2018) and Census data on TV ownership. Moreover, I collect price information from TV stations' "rate cards" and compute the revenue of local shows. TV shows provided an enormous step-up in the revenue and audience of entertainment shows.[26] Live shows reached on average 1,165 people before TV, while the median TV station could reach around 75,000 households. To track the group of workers who may

[22]I normalize by the average number of entertainers in treated labor markets to simplify interpretation of the regression coefficients as percentage point changes. Results without the normalization, as presented in Appendix 11, are in line with the baseline.

[23]For top income shares I focus on the larger 350 markets; details are described in Appendix 10.3.

[24]Note that the definition of mobility varies across Census vintages. Moreover, it does not distinguish between moves within and across labor markets. Noise in the outcome variable will inflate standard errors but not necessarily bias the estimates.

[25]The theatrical guide covers seating capacity and ticket prices of over 3,000 performance venues that cover ca. 80% of US local labor markets. According to the author "Information has been sought from every source obtainable - even from the Mayors of each of the cities." Undoubtedly the coverage will be imperfect and small or pop-up venues will be missed. Since we focus on star venues this omission may be of lesser concern. I use the largest available audience in the labor market as proxy for stars' show audience. I probe the reliability by manually comparing specific records with information from archival data and the data seems reliable.

[26]Details on revenue data are in Appendix 10.3. For TV shows, prices are imputed based on an demand elasticity estimated in a subset of 451 markets where data is available.

lose out from TV, I collect information on attendance and spending at county fairs. The data span revenues and ticket sales for over 4,000 county fairs over 11 years (1946-1957) and covers the majority of US labor markets. I collect these records from copies of the "Cavalcade of Fairs," which is published annually as a supplement to Billboard magazine and reports detailed records on county fairs. I aggregate local spending at the regional level and in three spending categories that are differentially close substitutes for television: spending on live shows (e.g., grandstand shows), fair tickets, and carnival items (e.g., candy sales and fair rides). Fair shows most closely resembled TV shows at the time, while candy sales and fair rides are by nature less substitutable with TV. Finally, I trace when county fairs faced competition from TV shows, using data on TV signal from Fenton and Koenig (2018).[27] The availability of TV in 1950 is shown in Figure 3, due to the freeze a number of places had narrowly missed out on TV signal at this point and the figure illustrates this variation too.

4 Empirical Results

During the roll-out of TV inequality in entertainer wages increased sharply. The sector started out far more equal than it is today. Figure 4a shows that before the advent of TV, most entertainers earned close to average pay, while after the introduction of TV pay dispersion grew substantially. Between 1939 and 1969 compensation of top entertainers grew disproportionately, many mid-income jobs had disappeared and a larger low-paid sector had emerged. At the same time, employment in performance entertainment remained stable, while it grew quickly in other leisure related activities, e.g. restaurant and bar workers, fountain workers and sport instructors (Figure 4b).This pattern of rising dispersion in log pay and lack of employment growth is precisely what superstar effects predict. Yet from these aggregate patterns it is unclear whether the simultaneous rise of TV is just a coincidence or causing these effects.

An ideal test of the superstar effects would randomize production technologies across labor markets. To get close to this ideal, I exploit the staggered introduction of television across local labor markets (m) over time (t) in a difference-in-differences regression:

$$Y_{mot} = \alpha_m + \delta_{ot} + \gamma X_{mt} + \beta TV_{mt} + \epsilon_{mot} \quad (6)$$

[27]Similar TV signal data has been widely used to study the effect of TV watching (e.g. Gentzkow, 2006; Gentzkow and Shapiro, 2008).

where α_m and δ_{ot} are labor market and occupation-year fixed effects; X_{mt} is a vector of time varying labor market characteristics. The treatment variable, TV_{mt}, uses the newly collected data on the number of local TV stations producing local shows in a market m at time t. I run the regression at the more disaggregated labor market (m), year (t), occupation (o) level to control for potential time fluctuations in the occupation definition with occupation-year fixed effects. The standard errors $\epsilon_{m,o,t}$ are clustered at the local labor market level, so that running the analysis at the disaggregated level will not artificially lower standard errors.

The variation in TV_{mt} comes from the staggered deployment of TV stations (See Section 3 for details). By leveraging regional differences in exposure to technical change, the identification can hold aggregate effects fixed and thus address trends in regulation and norms.

4.1 Results: Rising Returns at the Top

First, I test the headline prediction of the Superstar Effect, a sharp increase in wages at the top of the distribution (see Proposition 2.1a). I implement this test in the difference-in-differences analysis in equation 6 and test the response of the top percentile of entertainer wages to the launch of TV filming. This compares 546 local labor markets over time, excluding labor markets that are too small to compute the top percentile. The results strongly confirm this headline prediction and show a large and significant wage increase at the top. The top wage percentile increases by 14 log points or approximately 15% (see panel A in Table 1). Next, I check that this effect is indeed specific to the entertainment sector and study the position of entertainers in the US wage distribution. For this I compute the share of entertainers in the top percentile of the US wage distribution, which is available for the full sample and compares 722 local labor markets across the mainland US. Panel B in Table 1 shows that the share of top paid entertainers grows by 4 percentages points and thus nearly doubles. The results thus confirm that the launch of a TV station leads to a large and significant increase in top incomes in the entertainment sector, both in absolute and relative terms.[28]

Superstar effects predicts that the biggest gains from an expansion in market

[28] One potential concern with the latter approach are fluctuations in the denominator which could bias the estimates. Recall however that we constructed the variable with this challenge in mind and are holding the denominator fixed. A related approach uses per capita shares which suffer less from fluctuations in the denominator. Such estimates yield similar results (panel C of Table 1).

reach accrue to the most sought-after performers, and I turn to this prediction next. A first challenge for such a test is to get data on who is a popular performer before the launch of TV. I compute a measure for willingness to pay for entertainers based on the wage information in the de-anonymized Census data in 1939. I then link this variable to a measure of later TV success based on the "Who is Who of Television." One point in case is "uncle Ed" of Kentucky, who became a local TV star at WAVE TV. I find this local TV celebrity in the 1940 Census and identify his wage before TV was around. The linked data are a small panel that combines data on the popular appeal of entertainers before TV to their later TV success. Figure 5 shows the earnings percentiles of TV stars prior to the launch of TV. This reveals that the vast majority of TV stars were in the top tail of the entertainer wage distribution even before TV. The successful TV stars are thus the stars of the pre-TV era, and TV indeed amplifies the most successful entertainers.

Probing the Identification Assumption

The key identification assumption is that TV launch dates are unrelated to local trends. One piece of support for this assumption is that government documents show that locations are selected based on pre-determined local characteristics and the roll-out thus does not respond to local economic conditions. The 1952 "Final Allocation Report" for instance prioritizes locations by their local population in 1950. We can condition on such pre-determined local characteristics and then exploit the fact that approvals arise quasi-randomly from administrators working through the stack of priorities over time. To further verify that the timing of approval is unrelated to local trends I first control directly for time-varying changes in local labor markets. I run two specifications, one controlling for time varying local characteristics and one that allows for local labor market specific trends (column 2 and 3 in Table 1). The second approach adds more than 700 additional location specific trends and thus is a very demanding specification, standard errors increase accordingly. Both specifications find effects very similar to the baseline, indicating that differential local trends are not driving the findings.

To probe this assumption further, I run a placebo test with TV launches that are unexpectedly blocked. In addition to previous tests, this allows me to test if there are demand shocks that coincide with the launch of television. Recall that the interruption was a blanket freeze that put all license procedures on hold. For identification this indiscriminate approach is useful as it generates variation that is independent of local economic conditions. Figure 6 shows the impact of the freeze on TV launches and shows the sudden halt in approvals. I use this sudden stop and test for spurious effects in labor markets that narrowly miss out on TV. This placebo test compares blocked to untreated labor markets in a dynamic difference in differences regression with blocked stations ($TV_{mt}^{blocked}$) as treatment:

$$Y_{mot} = \alpha_m + \delta_{ot} + \gamma X_{mt} + \sum_t \beta_t TV_{mt}^{blocked} + \epsilon_{mot} \tag{7}$$

The effect of blocked stations is captured by β_t and is depicted in Figure 7a. The results show no effect in blocked locations and thus confirm that trends in blocked and untreated areas evolve in parallel. The TV roll-out therefore appears to unfold orthogonally to local demand shocks, shoring up confidence in the identification assumption of the above results. This test is arguably more convincing than a pre-trend test or a test based on placebo occupations, as we can test for local shocks in the same occupations and year. For completeness, I perform those additional robustness checks and they show the same result (Appendix 10.2.3).

Finally, I show that the treatment effects only last while local filming is the predominant practice. With the advent of the videotape, when national filming takes over, the local effects disappear (Figure 7b). I capture the effect on the newly emerging hubs, by including a pre-determined proxy for local filming cost, interacted with the period after the launch of the videotape. In this period without location constraints on filming, places with a comparative advantage in filming see fast top income growth. The remainder of previous filming centers, however, experience a striking decline and the difference between such places and the control areas disappears. This rise and subsequent disappearance of the effects confirms that the results are driven by TV, rather than by diverging local trends. In 1969 the differences between treatment and control group reverted to the pre-treatment level, suggesting that the common trend assumption holds.[29]

Links Between Markets

So far the analysis focused on changes in inequality. For the interpretation of the results, it will be useful to distinguish between two potential mechanisms: migration of entertainers and changing returns to talent. I therefore directly examine the

[29] Notice that this test is a powerful parallel trend check that leverages both pre and post TV periods for a parallel trend check. A conventional pre-trend check is reported in Appendix 10.2.2.

mobility of entertainers across local labor markets.[30] Note that labor markets where entertainers could reach the largest audiences before TV also tend to receive TV earlier. Such differences will be absorbed by the location fixed effects and, absent mobility frictions, we would not expect substantial mobility responses. This frictionless benchmark may however not hold in practice, so I test the mobility response empirically. The results indeed show very modest changes in mobility patterns. The point estimates are in fact negative and confidence intervals are tight and rule out that mobility increased by more than 2% (columns 1 – 3 of Table 2). Using these results to bound the impact of migration, such effects can explain at most a quarter of the total effect, while the central estimates suggest that mobility plays next to no role in explaining the results. A related test studies mobility across neighboring labor markets where moving is arguably easiest. Excluding neighbors of treated areas should thus alleviate mobility effects. Results that exclude neighboring areas are close to the baseline, indicating again that migration plays a minor role for the findings (panel B of Table 2). This suggests that the results mainly capture changes in labor market returns.

4.2 Distinguishing the Superstar Mechanism

4.2.1 Results: Demand for Non-Stars

Superstar effects move labor market closer to a winner-takes-all market and in the process leads to falling employment in the industry (proposition 2.1 d)). To test the employment effects of TV, I study entertainer labor markets where customers can access TV entertainment and compare entertainer employment in local labor markets with differential access to TV signal. Such regressions thus rely on variation in TV signal rather than TV filming that was relevant above. The TV signal data is not available at a channel level, instead I use a dummy for access to TV signal as regressor. Employment records in the Census are available for additional years, which allows me to expand the sample period backward by a decade to 1930. Results for this extended period are reported alongside results for the baseline period. I find that the introduction of TV leads to a substantial fall in local entertainer employment and I estimate that around 13% of jobs are lost (Table 3 column 1, panel A for the

[30] Note that panel data, albeit not available, would not separately identify the effect of mobility and changes in productivity. In the Superstar model mobility affects the assignment of workers to firms and an individual fixed effect regression would conflate time varying changes in matching with changes in productivity.

extended sample and panel B for the baseline sample). This confirms the prediction of superstar effects and is sharply at odds with models where technical change causes a positive demand shock, which would raise employment.

Since these specifications use variation from TV signal rather than from TV filming, it is salient to probe the identifying assumption again. As before, results are robust to the inclusion of controls and local trends (Columns 2 and 3 of Table 3).[31] Further, common trend tests suggest that the set-up is valid. A first pre-trend test introduces a lead to the treatment variable in the difference in differences regression, which captures differential changes in treatment and control areas, right before the treatment and shows no sign of such differential trends (Column 4).[32] Placebo tests with stations that did not happen due to the freeze also show no effect (Panel C of Table 3), adding confidence that there are no spurious trends during the roll-out.

The shift of demand from a profession's mediocre workers towards its stars should also be reflected in a decline in mid-paid jobs (see Proposition 2.1c). To test for such effects, consider entertainers who are below the top 90th percentile of the US wage distribution but still in the upper quartile. These are entertainers who receive above-average pay but are far from the top of the entertainer distribution. TV has a significantly negative effect on this group, the number of jobs that pay in this range declines by around 50%. The results look similar between the median and the 75th percentile (results are reported in Figure 8). Television therefore leads to a substantial decline in well-paid jobs and makes it substantially worse to be a mediocre entertainer during the TV era.

The corollary to disappearing mid-paid jobs is the growing low-pay sector. Analyzing the share of entertainers paid below the median, we observe a modest rise in the share of entertainers with wages at the very bottom of the distribution, with little change in the second quartile. Television thus reduces the payoff of non-star talent and creates a growing low-pay sector.

4.2.2 Results: Fractal Inequality

A third implication of superstar effects is a widening wage gap between stars and their slightly less talented peers (see Proposition 2.1b). A non-parametric test of

[31] Median income is not available in 1930 and controls in the extended sample use the remaining variables.

[32] TV signal, unlike local filming, is not removed and we thus cannot rely on pre- and post-periods to identify counterfactual trends.

this prediction repeats the baseline difference-in-differences regression, focusing on percentiles just below the star level. Take, for example, entertainers who are below the top 1% but still among the top 5% of the US wage distribution. I find that television also benefits this group but the effect is only one tenth the size of the effect at the very top. Television therefore disproportionately benefits the superstars and widens the pay gap in the top tail of the distribution. To confirm this pattern we can look at the next lower wage bin, between the 90th and 95th percentile. Here the effect of television is insignificant, again confirming that television's effect fades quickly as we move away from the top stars in the market. The effect of technology declines remarkably quickly in the top tail. TV appearances help a small group of top stars, has moderate effects on backup stars and has no discernible benefit for other top earners.

The growing fractal inequality is also reflected in growing top income dispersion within entertainment. This is closely related to the previous results but focuses on an inequality measure widely used in the literature, top income shares.[33] Prior to TV, the fraction of income going to the 1% highest earners in a local labor market was, on average, 3.8%.[34] TV filming increased the top income share by 3.7 percentage points thus nearly doubling the income share (Table 4). Proposition 2.1b suggests that the growth in these shares should escalate toward the top of the distribution. Indeed, most of the gains in the top 1% accrued to the very highest earners. The top 0.1% of entertainers saw their income share rise by 2.4%. This group is only one tenth of the top 1% but accounts for over half of the rise of the top 1% income share. While the share of income going to the top 1% doubled and the equivalent share of the top 0.1% grew 4 fold, the top 10% share grew only 30%. A test of equal growth rates in the top tail is strongly rejected, which aligns with superstar effects where wage growth is strongest at the very top of the distribution.[35]

In summary, the effects across the entire distribution of entertainer pay are U-shaped (Figure 8). At the very top of the distribution TV has a large positive impact, but such positive effects decline quickly as we move away from the very top, turning negative below the 90th percentile. At the same time we see a growing low-paid sector in the industry. This characteristic pattern offers direct empirical support for the Superstar Effect.

[33] See for example Piketty, 2014; Piketty and Saez, 2003.

[34] The equivalent number for the US economy as a whole is substantially larger at about 10%. This reflects that within a given region and industry income are less dispersed.

[35] The appendix confirms the results with a set of quantile regressions (see Appendix 10.2.4).

5 Magnitudes, Monopsony and the Labor Share

Finally, I study the magnitude of top income growth and the interaction of superstar effects with imperfect competition. To quantify those effects we need to express the observed change in terms of the elasticity of top pay with respect to the market size. This elasticity is also at the heart of previous applied work on superstar effects, Gabaix and Landier (2008), Terviö (2008) use this elasticity to calibrate the key structural parameters of the superstar model and then simulate top income growth among CEOs. The standard approach is to calibrate the elasticity to the correlation of top pay and market size, proxied by firm value. An attractive feature of the entertainment setting is that it offers a direct measure of workers' market reach, the entertainers' audience. Moreover, I can use plausibly exogenous variation in market reach to identify the parameter of interest. In other words, I supplement correlational estimates with an instrumental variables approach. These estimate use the following regression equation:

$$ln(w^{99}_{m,t}) = \alpha_0 + \alpha_1 ln(s^{99}_{m,t}) + \epsilon^{99}_{m,t} \tag{8}$$

where $w^{99}_{m,t}$ is the 99th percentile of the entertainer wage distribution in market m and year t, and $s^{99}_{m,t}$ is the size of the market that such entertainers can reach. Previous studies focused on variation over time or across industries to estimate α_1. The setting of this study by contrast studies variation within an industry across time and regions and can thus control for year and labor market fixed effects. Moreover, I can leverage the TV roll-out as an instrument for market reach and thus isolate plausibly exogenous variation in market reach.

The first stage of such an instrumental variables approach is the effect of TV on the audiences of star entertainers. This effect is in-itself of interest since superstar effects have sharp predictions about the change in audience after the launch of TV. Customers should shift from more mediocre entertainers to the stars of the profession and in turn increase the market value of stars. To study such shifts, I first look at the audience of shows of local stars. The launch of a television station increased the audience of the largest shows by about 150 log points. Converted to a growth rate this implies a growth of over 300%, or a fourfold increase in market size (Panel A of Table 5). It will be useful to also quantify the change in market size in dollar terms. Such estimates quantify the change in the value of a star performer that went hand in hand with the growth in audience size. In dollar terms market reach of stars roughly tripled (Panel B of Table 5). This confirms that the launch of TV stations dramatically increased the market value of top talent.

In the superstar model rising demand for stars is accompanied by declining interest in ordinary local live entertainment. To study such effects, I analyze data on county fairs, a form of entertainment widely available throughout the US. TV leads to a 5% decline in audiences and spending at local county fairs (column 1 and 2 of Table 6). These results are however noisy and hide substantial heterogeneity across types of entertainment. Splitting the results by types of entertainment, the data show substantial heterogeneity in the effects. Demand for entertainment that is similar to TV, such as grandstand shows, falls significantly, while demand for entertainment that is different from TV, e.g. candy sales and amusement rides, holds up (see columns 3 and 4 of Table 6 and panel B for regressions at the county level). This shows that even within entertainment, spending on close substitutes to TV are most affected by the availability of TV signal. These result confirm that TV reduced demand for local live entertainment, increased demand for star entertainers, and shifted marginal revenue productivity in favor of stars at the expense of ordinary entertainment.

Next, I turn to the elasticity of top pay to audience size and estimate equation 8. A large literature has studied the relation of pay and market size, usually proxied by firm size, and used cross-sectional regressions to estimate α_1. I replicate this approach with a cross-sectional OLS estimator in 1939 and leverage variation in the size of local theaters across local labor markets. In line with previous results I find a highly significant effect of market size on top pay with a point estimate for α_1 of 0.23 (see panel A. of Table 7). This means that moving from a local labor market with a small theater to a market twice the size increases pay for a top entertainer by 23%. The effect may of course reflect broader differences in local labor markets, beyond differences in market size. Indeed, after controlling for local labor market characteristics, the effect disappears almost entirely (column 2 of the same Table).[36]

The roll-out of TV allows me to compare such OLS estimates to an IV estimate. The first stage, the effect of TV on audience size, is is highly significant. The associated first stage F-statistic is around 20, well above conventional cutoffs. The IV estimator for α_1 is also highly significant with a point estimate of 0.17. This implies that wages at the 99th percentile grow 17% when market size doubles. While this wage effect is sizable, the effect is 30% lower than the cross-sectional OLS estimate

[36] The panel OLS estimate would compare wages across local labor markets over time, as market reach changes. However, in my data variation in market reach within a local labor market over time comes exclusively from the launch of TV and hence such a panel OLS is therefore mechanically close to the IV estimate.

above. This suggests that the correlation of market size and top pay is bigger than the causal effect of market size. One potential driver of this upward bias is the availability of better talent in bigger markets.

I next explore the role of monpsony power in labor markets. The benchmark superstar model is perfectly competitive and the models' predictions change sharply with imperfect labor market competition, as monopsony employers will not pass-on the surplus from greater scalability of production. This is of particular interest, since many modern scale related technologies appear to go hand in hand with a small number of companies that develop and control access to new scale related technologies. In the entertainment setting we can leverage government entry restrictions to test the role of competition for superstar effects. I allow for differential effects of TV in markets with a single TV station vs. markets with multiple TV stations. The differences between monopsonistic and competitive labor markets are striking. Markets with a monopsony employer see almost no top income growth, while gains are large when there are competing employers. These results are confirmed when I narrow in on the variation from the roll-out interruption (Table 8). These findings emphasize the importance of competition for superstar effects. Only when employers are competing for talent, does growing market scale translate into rising top pay. Finally, I study the implication for the labor share. A growing literature discusses the relation of rising superstars and the fall in the labor share. To link my results to this literature, I estimate how rising market value of top talent is distributed. This is implemented by running regression 8 with data on revenues as endogenous regressor and TV as instrument.[37] TV is a strong instrument with an F-statistic between 28 and 57. The 2SLS estimate shows that one dollar growth in market value of top talent leads to 22 cents higher pay for star workers.[38] A constant labor share would require that pay grows proportionally to revenues, aka an elasticity of 1.[39] These results thus strongly confirm that rising superstar effects go hand in hand with a falling labor share.

[37]A drawback of focusing on revenues is that revenues conflate quantity and endogenous price effects. The response of revenues thus doesn't naturally map into map into model predictions.

[38]Note that this estimate is bigger than the elasticity with respect to audience size. This difference arises because the launch of TV reduced the cost of top entertainment for consumers, which implies that the first stage effect on revenues is relatively smaller than the one on audience size. The smaller first-stage effect increases the IV estimate.

[39]Estimates of this elasticity among CEOs range between 0.1 and 1, my IV estimate thus falls into the lower half of this range (Frydman and Saks, 2010; Gabaix and Landier, 2008).

6 Conclusion

Little is known about the causes of the vast changes in top incomes observed in recent decades. Superstar Effects link these changes to technical innovation, particularly in communication technologies, where it is easier to operate over distances. This paper provides causal evidence on the effect of growing production scalability on wages and provides an empirical test of the Superstar Effect.

To test the Superstar Effect, I exploit quasi-experimental variation in market reach in the entertainment industry and show that the staggered introduction of TV substantially changed audience sizes for entertainment shows. Star entertainers increased their audiences fourfold through TV, and the sector experienced sharp income concentration at the top. The increase in production scalability has profound effects on inequality at both the top and bottom of the distribution, in line with the prediction of superstar effects. The characteristic patterns of superstar effects are strongly supported in the data. Income growth escalates as we move up towards the top of the wage distribution and the share of income going to the top 1% nearly doubles. Moreover, the ability to reach larger markets puts many lesser stars out of work. The number of mid-paid entertainer jobs declined significantly and total employment fell about 13%.

The paper also finds that competition for talent is a key driver of superstar effects, while top income growth is muted in settings with limited competition in the labor market. This highlights that market concentration on a few stars does not necessarily indicate malfunctioning of markets, instead the superstar effects suggests that rising market concentration is a sign of technical progress. To evaluate inefficiencies associated with top income concentration, it will be important to distinguish cases where superstar effects bring better quality to a greater share of consumers from cases where market concentration results from the break-down of competition.

Further research is also needed to assess the magnitude of superstar effects for overall top income growth. It is important to keep in mind that superstar effects arise when production becomes scalable and talent is heterogeneous and unique. Settings where talents are close substitutes likely exhibit smaller superstar effects. As a result, the estimates of the entertainment sector likely provide an upper bound for superstar effects in the aggregate economy, and further research is needed to measure superstar effects in additional sectors.

References

Acemoglu, D. (1998). Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality. *The Quarterly Journal of Economics*, *113*(4), 1055–1089.

Acemoglu, D., & Restrepo, P. (2019). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy*.

Akerman, A., Gaarder, I., & Mogstad, M. (2013). The Skill Complementarity of Broadband Internet. *Quarterly Journal of Economics*, (7762).

Autor, D. H., & Dorn, D. (2013). The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market. *American Economic Review*, *103*(5), 1553–97.

Bakija, J., Cole, A., & Heim, B. (2012). Jobs and Income Growth of Top Earners and the Cause of Changing Income Inequality: Evidence from US Tax Return Data. *unpublished manuscript, Williams College*.

Bartel, A., Ichniowski, C., & Shaw, K. (2007). How Does Information Technology Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement and Worker Skills. *The Quarterly Journal of Economics*, *122*(November), 1721–1758.

Bound, J., & Johnson, G. (1992). Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations. *America Economic Review*, *82*(3), 371–392.

Brynjolfsson, E., & McAfee, A. (2011). *The Race against the machine*.

Card, D., & DiNardo, J. (2002). Skill Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles. *Journal of Labor Economics*, *20*(4).

Caselli, F., & Manning, A. (2019). Robot Arithmetic: New Technology And Wages. *CEP Discussion Paper*.

Célérier, C., & Vallée, B. (2019). Returns to Talent and the Finance Wage Premium. *The Review of Financial Studies*, *32*(10), 4005–4040.

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *Quarterly Journal of Economics*, *129*(4), 1553–1623.

Chetverikov, D., Larsen, B., & Palmer, C. (2016). IV Quantile Regression for Group-Level Treatments, With an Application to the Distributional Effects of Trade. *Econometrica*, *84*(2), 809–833.

Eckert, F., Ganapati, S., & Walsh, C. (2019). Skilled Tradable Services: The Transformation of U.S. High-Skill Labor Markets. *Working Paper*.

Fenton, G., & Koenig, F. (2018). The Labor Supply Response to Innovation in Entertainment: Evidence from TV. *Working paper*.

Frydman, C., & Saks, R. E. (2010). Executive compensation: A new view from a long-term perspective, 1936-2005. *Review of Financial Studies*, *23*(5), 2099–2138.

Gabaix, X., & Landier, A. (2008). Why Has CEO Pay Increased So Much? *Quarterly Journal of Economics*, *123*(1), 49–100.

Garicano, L., & Hubbard, T. N. (2009). Specialization, firms, and markets: The division of labor within and between law firms. *Journal of Law, Economics, and Organization*, *25*(2), 339–371.

Gentzkow, M. (2006). Television and voter turnout. *The Quarterly Journal of Economics*, *121*(August), 931–971.

Gentzkow, M., & Shapiro, J. (2008). Preschool Television Viewing and Adolescent Test Scores: Historical Evidence from the Coleman Study. *Quarterly Journal of Economics*, *123*(February), 279–323.

Goos, M., Manning, A., & Salomons, A. (2014). Explaining Job Polarization: Routinization and Offshoring. *American Economic Review*, *104*(8), 2509–26.

Guellec, D., & Paunov, C. (2017). Digital Innovation and the Distribution of Income.

Kaplan, S. N., & Rauh, J. (2010). Wall street and main street: What contributes to the rise in the highest incomes. *Review of Financial Studies*, *23*(3), 1004–1050.

Kaplan, S. N., & Rauh, J. (2013). It's the Market: The Broad-Based Rise in the Return to Top Talent. *Journal of Economic Perspectives*, *27*(3), 35–56.

Katz, L. F., & Murphy, K. M. (1992). Changes in Relative Wages, 1963-1987: Supply and Demand Factors. *The Quarterly Journal of Economics*, *107*(1), 35–78.

Krueger, A. B. (2005). The Economics of Real Superstars: The Market for Rock Concerts in the Material World. *Journal of Labor Economics*, *23*(1), 1–30.

Lemieux, T. (2006). Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill? *American Economic Review*, *96*(3), 461–498.

Michaels, G., & Graetz, G. (2018). Robots at Work. *Review of Economics and Statistics*.

Piketty, T. (2014). *Capital in the 21st Century*. Harvard University Press.

Piketty, T., & Saez, E. (2003). Income Inequality in the United States, 1913-1998. *The Quarterly Journal of Economics*, *68*(February), 1–39.

Rosen, S. (1981). The Economics of Superstars. *American Economic Review*, *71*(5), 845–858.

Sattinger, M. (1975). Comparative advantage and the distributions of earnings and abilities. *Econometrica*, *43*(3), 455–468.

Scheuer, F., & Werning, I. (2017). The Taxation of Superstars. *Quarterly Journal of Economics*, *132(1)*, 211–270.

Sterne, J. (1999). Television under construction: American television and the problem of distribution, 1926-62. *Media, Culture & Society*, *21*(199905), 503–530.

Terviö, M. (2008). The difference that CEOs make: An assignment model approach. *American Economic Review*, *98*(3), 642–668.

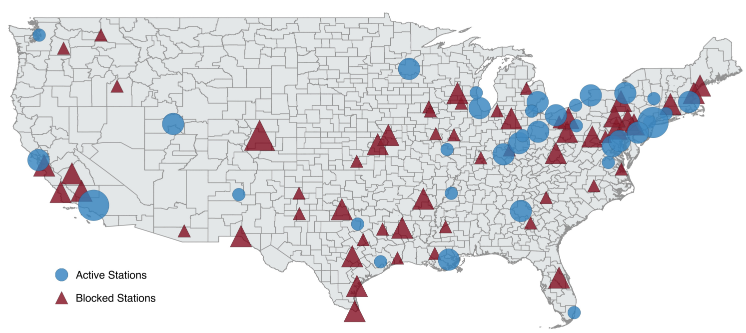
Tinbergen, J. (1956). On the Theory of Income Distribution. *Weltwirtschaftliches Archiv*, *77*(2), 155–175.

7 Figures

Figure 1: Superstar Effect: Employment Change at Different Wage Levels

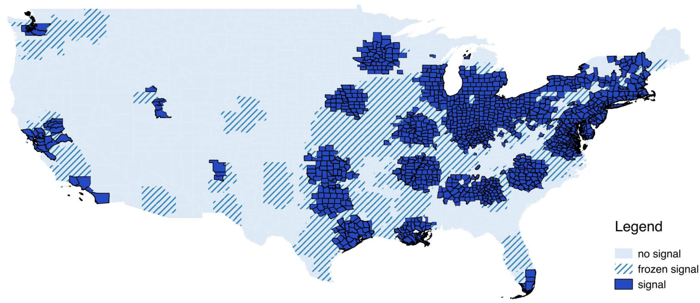
[Notes] The figure shows the impact of superstar effects on employment changes across the wage distribution. The figure is based on equation 5 for parameterization $g_e = 0.2x^{(1.3)} - 1$. The figure reports growth rates across the full distribution by grouping job growth outside the range of previous support with the final bins with positive mass and thus avoids dividing by zero.

Figure 2: TV Filming of Licensed and Blocked Stations in 1949



[Notes] Symbols show the location of TV filming and the size of a symbol indicates the number of TV stations per local labor market. Active stations are blue circles, frozen stations red triangles. Source: FCC reports.

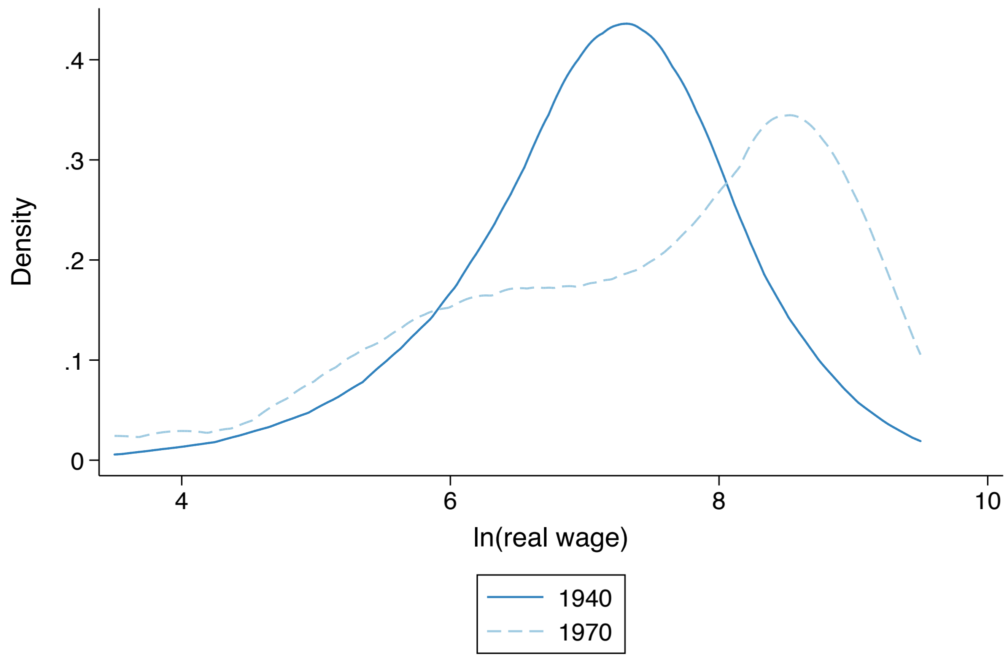
Figure 3: TV Signal of Licensed and Blocked Stations in 1949



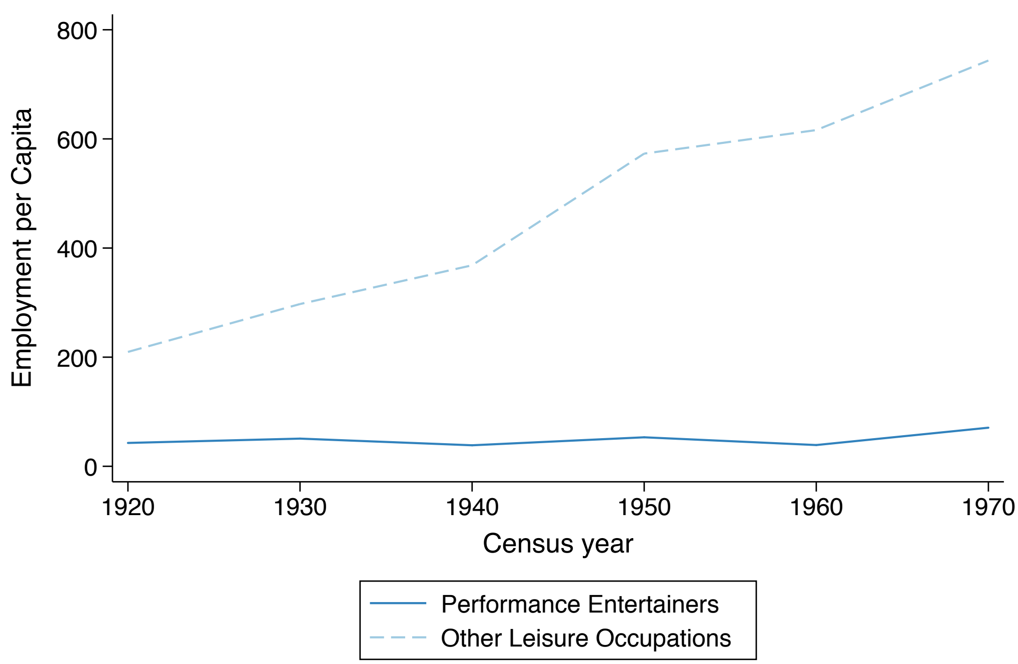
[Note] Areas in dark blue can watch TV, while shaded areas would have had TV signal from blocked TV stations. Signal coverage is calculated using an Irregular Terrain Model (ITM). Technical station data from FCC files, as reported in TV Digest and Television yearbooks, are fed into the model . Signal is defined by a signal threshold of -50 of coverage at 90% of the time at 90% of receivers at the county centroid. Source: Fenton and Koenig (2018).

Figure 4: Change in Entertainment 1940 – 1970

(a) Entertainer Wage Distribution

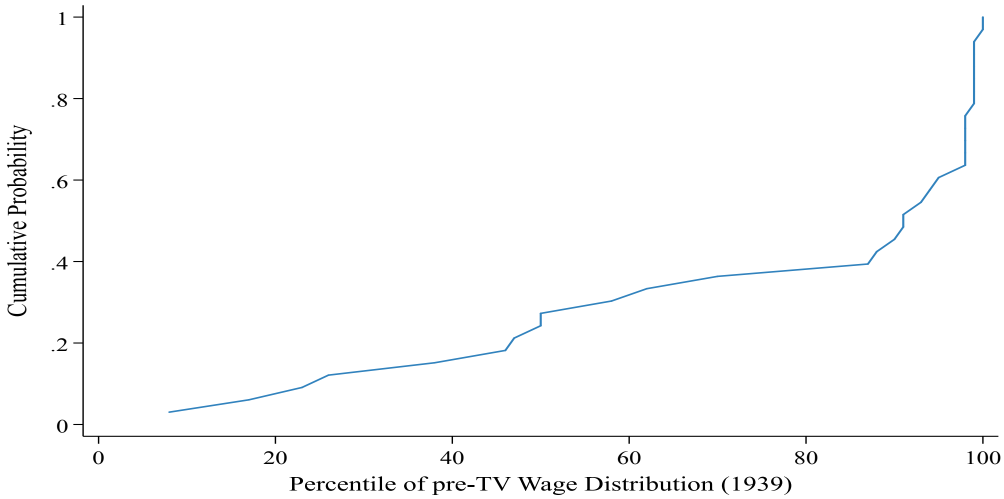


(b) Entertainer per Capita



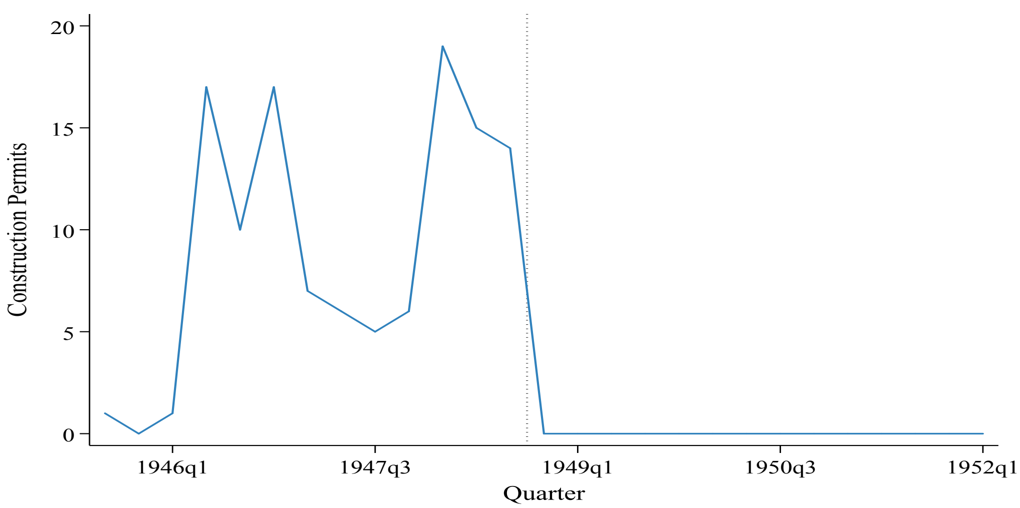
[Notes] Panel A shows the entertainemnet log real wage distribution in 1940 and 1970 from the lower 48 states. Dollar values are in 1950 USD. Density is estimated using the Epanechnikov smoothing kernel with a bandwidth of 0.4 and Census sample weights. Common top code applied at $85,000. Panel B shows employment per 100,000 inhabitants of performance entertainers (defined in text) and other leisure related occupations (drink & dine and "other entertainment occupations"). The mean for performance entertainers is 49 and for other leisure occupations 468. Sources: US Census 1940, 1970.

Figure 5: Position of Future TV Stars in the 1939 US Wage Distribution



[Note] The Figure shows the CDF of wage distribution ranks of TV stars before they became TV stars. TV stars are defined in the 1950 "Who is Who of TV". These individuals are linked to their 1939 Census wage records. 1939 wages are corrected for age, education and gender using a regression of log wages on a cubic in age, 12 education dummies and a gender indicator. Source: Radio Annual, Television Yearbook 1950.

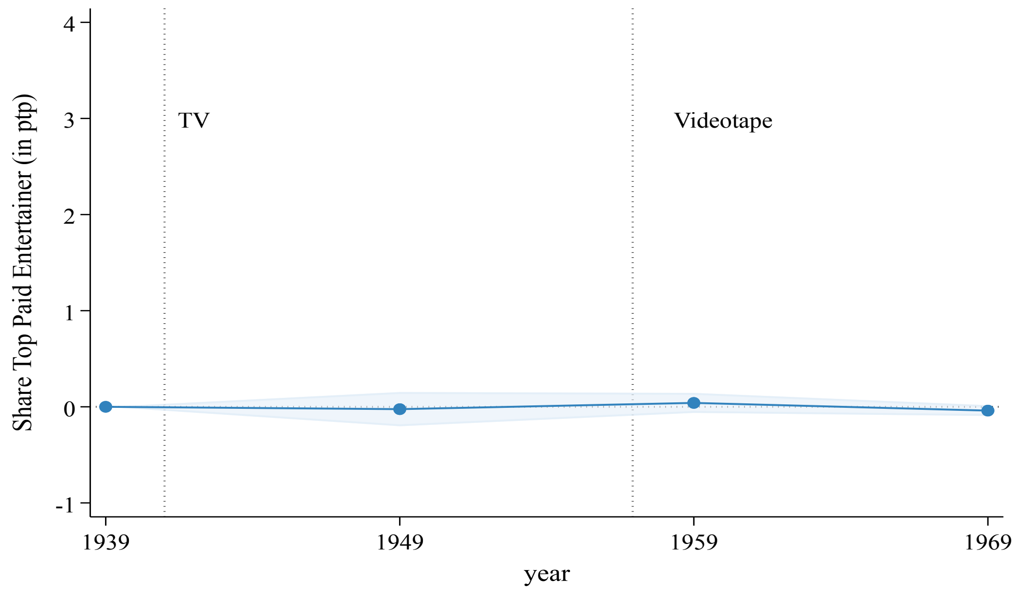
Figure 6: Number of TV Licenses Granted



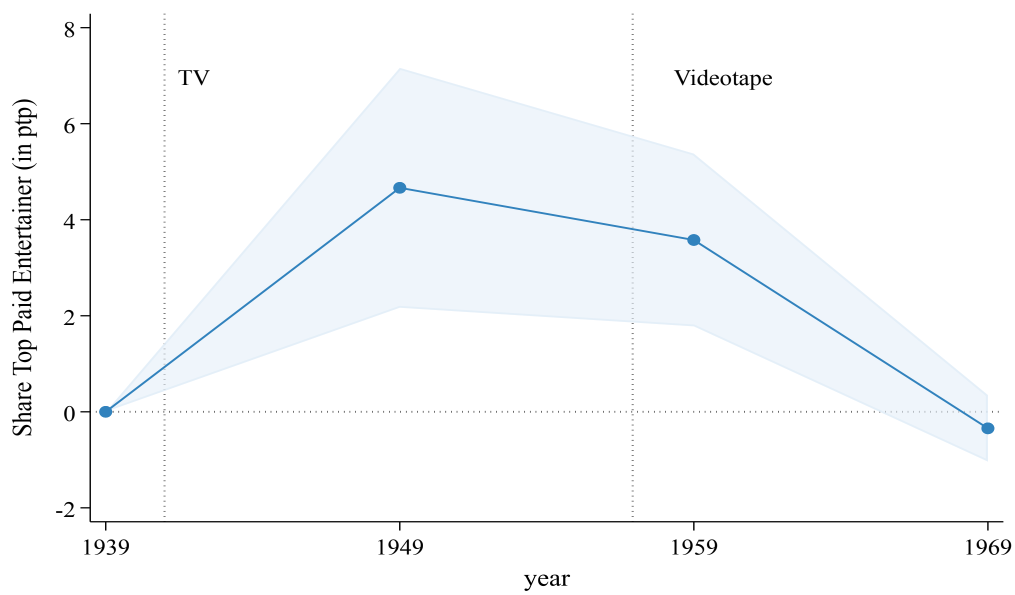
[Note] Missing issue dates of construction permits are inferred from start of operation dates. Source: TV Digest reports.

Figure 7: Dynamic Treatment Effect of TV on

(a) *Blocked TV Stations*

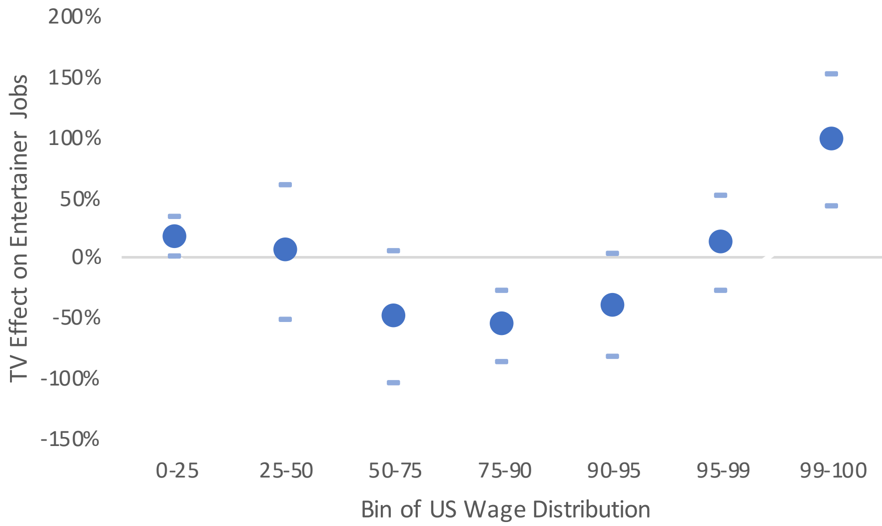


(b) *Active TV stations*



[Note] Figure plots treatment coefficients from two dynamic difference in differences regressions. Panel a) shows the coefficient on $FrozenTV_{m,t}$ (comparison groups are untreated areas) and panel b) shows the coefficient on $TV_{m,t}$. Top-paid entertainers are in the top 1% of the US income distribution. Vertical lines mark the beginning of local TV ("TV") and the end of local TV ("Videotape"). The area shaded in light blue marks the 95% confidence interval. Standard errors are clustered at the local labor market level.

Figure 8: Effect of TV on Entertainer Employment Growth at Different Wage Levels



[Note] Each dot is the treatment effect estimate of a separate DiD regression. It shows a TV station's effect on entertainer jobs at different parts of the wage distribution. Percentile bins are defined in the overall US wage distribution. Dashes indicate 95% confidence intervals. See table 1 for details on the specification. Sources: US Census: 1940-1970.

8 Tables

Table 1: Effect of TV on Entertainer Top Earners

	Panel A: 99^{th} Percentile of Entertainer Wages (log)		
Local TV stations	0.138	0.126	0.100
	(0.030)	(0.031)	(0.042)
Effect/Baseline	14.8%	13.4%	10.5%
CZs (Cluster)	541	541	541
	Panel B: Entertainer among Top 1% of US Earners (% of Entertainers)		
Local TV Stations	4.14	4.31	5.93
	(1.26)	(1.27)	(2.21)
Effect/Baseline	92%	96%	132%
CZs (Cluster)	722	722	722
	Panel C: Entertainer among Top 1% of US Earners (Per Capita)		
Local TV Stations	0.40	0.40	0.31
	(0.10)	(0.10)	(0.10)
Effect/Baseline	133%	133%	103%
CZs (Cluster)	722	722	722
Time & CZ FE	Yes	Yes	Yes
Demographics	–	Yes	–
Local labor market trends	–	–	Yes

[Note] Outcomes: Panel A: The entertainer wage at the 99th percentile, Panel B: share top-paid entertainers, Panel C: top-paid entertainer per capita in 10,000s. Specifications: Each cell is the result of a separate DiD regression on the number of TV stations in the local area. All regressions control for CZ & time fixed effects and local filming cost in years after the invention of the videotape . Demographics: median age & income, % female, % black, population density and trends for urban areas; local labor market trends: allow for a linear trend for each local labor market. Entertainers are actors, athletes, dancers, entertainers nec, musicians. Panel A uses the quantile DiD estimator developed by Chetverikov, Larsen, and Palmer (2016); cells where the 99th percentile cannot be computed are dropped. The unit of analysis is the CZ – year level in A and the the more disaggregated CZ – occupation – year level in B and C to additionally control for year-occupation fixed effects. Panel A uses 1,435 observations and Panel B and C 13,718 observations, demographic data is missing for one CZ in 1940. "Effect/Baseline" reports treatment effects relative to the baseline value of the outcome variable. Observations are weighted by local labor market population. Standard errors are reported in brackets and are clustered at the local labor market level. Sources: US Census 1940-1970.

Table 2: Effect of TV on Mobility Between Labor Markets

	Panel A: Share Entertainers who Migrated		
Local TV stations	-0.014	-0.017	-0.010
	(0.015)	(0.015)	(0.020)
	Panel B: Entertainer among Top 1% of US Earners (excl. neighbor)		
Local TV stations	4.30	4.46	6.16
	(1.31)	(1.30)	(2.27)
Time-Occupation & CZ FE	Yes	Yes	Yes
Demographics	–	Yes	–
Local labor market trends	–	–	Yes

[Note] Dependent variables are, Panel A the fraction of entertainers who moved, Panel B share of Entertainers among the top 1% of the US wage distribution, excluding labor markets that neighbor treated labor markets. Specification details are as in Table 1, except that Panel B is run on a reduced sample of 10,792 observations. Source: US Census 1940-1970.

Table 3: Effect of TV on Entertainer Employment

	Ln(Employment in Entertainment)			
	(1)	(2)	(3)	(4)
	Panel A: TV Signal 1930-1970			
TV signal$_{t+1}$				0.039
				(0.033)
TV signal$_t$	-0.133	-0.127	-0.125	-0.123
	(0.059)	(0.059)	(0.061)	(0.060)
	Panel B: TV Signal 1940-1970			
TV signal$_t$	-0.128	-0.114	-0.134	
	(0.061)	(0.061)	(0.063)	
	Panel C: Placebo TV Signal			
Placebo TV signal$_t$	0.053	0.044	0.053	
	(0.083)	(0.083)	(0.084)	
Clusters	722	722	722	722
Time-Occupation & CZ FE	Yes	Yes	Yes	Yes
Demographics	-	Yes	-	-
Local Labor Market Trends	-	-	Yes	-

[Note] Dependent variable "ln(Employment in Entertainment)" is the inverse hyperbolic sine of employment in entertainment. Control variables and specifications are as described in Table 1, except that demographic controls exclude median income. TV signal is a dummy that takes value 1 if signal is available in a commuting zone. Placebo TV signal is the signal of stations that were blocked. Subscript "t+1" refers to the lead of the treatment. Standard errors, reported in brackets, are clustered at the local labor market level. Source: TV signal from Fenton and Koenig (2018) and labor market data from US Census 1930-1970.

Table 4: Effect of TV on Top Income Shares in Entertainment

	Share of Income		
	Top 0.1%	Top 1%	Top 10%
Local TV stations	2.37	3.71	6.08
	(1.27)	(1.69)	(2.12)
Time & CZ FE	Yes	Yes	Yes
Effect/Baseline	239%	96%	33%
P-value: same growth as top 1% share	0.0043	—	0.0000

[Note] Dependent variable top p% is the share of income going to the top p percent of entertainers in a given local labor market-year. The shares are calculated using Pareto interpolation as described in the text. The sample includes the larger 350 labor markets and 1,069 observations. Estimates are based on a difference in difference specification. P-values from a test of equal growth rates in top income shares are also reported. This test is implemented in a regression with the ratio of top income shares as dependent variable. Standard errors are clustered at the local labor market level. Sources: US Census 1940-1970.

Table 5: Effect of TV on Market Reach of Local Stars

	(1)	(2)	(3)
		Panel A: Show Audience (log)	
Local TV stations	1.499	1.526	1.146
	(0.240)	(0.223)	(0.220)
Effect/Baseline	348%	360%	215%
		Panel B: Show Revenue (log)	
Local TV stations	1.095	1.116	1.146
	(0.207)	(0.168)	(0.220)
Effect/Baseline	199%	205%	215%
Clusters	722	722	722
Time & CZ FE	Yes	Yes	Yes
Demographics	–	Yes	-
Local labor market trends	–	–	Yes

[Note] Dependent variables are, Panel A: potential show audience of the largest show in the commuting zone, computed from venue seating capacity and TV households in transmission area, Panel B: potential revenue of largest show. Cells report results from separate DiD regressions across local labor markets. Control variables are as described in Table 1. The total number of CZ - year observations are 2,656. Sources: See text.

Table 6: Effect of TV on Spending at Local County Fairs

	(1) Fair Visits (log)	(2) Ticket Receipts (log)	(3) Show Receipts (log)	(4) Carnival Receipts (log)
	Panel A: Local Labor Market Level			
TV signal	-0.051	-0.047	-0.059	0.014
	(0.031)	(0.024)	(0.022)	(0.022)
Clusters	722	722	722	722
Time & Labor Market FE	Yes	Yes	Yes	Yes
	Panel B: County Level			
TV signal	-0.013	-0.014	-0.018	0.001
	(0.010)	(0.007)	(0.007)	(0.006)
Clusters	3,111	3,111	3,111	3,111
Time & County FE	Yes	Yes	Yes	Yes

[Note] Dependent variables are summed across county fairs in location m in year t at annual frequency from 1946 to 1957. All variables use the the inverse hyperbolic sine transformation to approximate the log function, while preserving 0s and monetary variables are in 1945 US Dollars. In Panel A the unit of observation m is a local labor market and in Panel B a county. Treatment is the number of TV stations that can be watched in the commuting zone. Data on carnival receipts (col 4) are unavailable for 1953 and 1955. Panel A uses 8,664 local labor market observations (7,220 in column 4), while Panel B uses 37,332 county observations (in col 4 31,110). Standard errors, reported in brackets, are clustered at the local labor market level in Panel A and at the county level in Panel B. Source: Billboard Cavalcade of Fairs 1946-1957 and Fenton and Koenig (2018).

Table 7: Elasticity of Entertainer Top Pay to Market Reach

	(1)	(2)	(3)
	99^{th} *Percentile of Entertainer Wages (log)*		
		Panel A: OLS - Cross-section 1939	
ln(Audience size)	0.234	0.023	
	(0.036)	(0.036)	
		Panel B: IV	
ln(Audience size)	0.166	0.149	0.149
	(0.017)	(0.019)	(0.024)
First-stage F-statistic	33.3	25.7	20.0
		Panel C: IV	
ln(Value of market ($))	0.220	0.192	0.198
	(0.028)	(0.022)	(0.036)
First-stage F-statistic	57.10	38.1	28.7
Demographics	–	Yes	–
Local labor market trends	–	–	Yes

[Note] Dependent variable is the entertainer wage at the 99th percentile. Panel A reports coefficients from a cross-sectional regression that uses variation across 573 local labor markets in 1939. Panel B and C show results from an IV regression that uses TV stations as instrument and uses the full panel with 2,148 observations. The corresponding first stage and reduced form results are reported in table 1 and table 5. The first-stage F-statistic is the Kleibergen-Paap F-statistic that allows for non-iid standard errors. Control variables are described in table 1 and market reach measures in table 5. Standard errors are clustered at the local labor market level. Sources: see table 1 and table 5.

Table 8: Effect of Competition in Labor Markets

	(1)	(2)	(3)
	Entertainer in US Top 1%		
Local TV station (dummy)	5.90	0.753	-0.57
	(3.06)	(1.91)	(0.36)
Multiple local TV station (dummy)		9.07	10.37
		(4.99)	(4.70)
Frozen competitor			1.43
			(2.10)
Clusters	722	722	722
Time-Occupation & CZ FE	Yes	Yes	Yes

[Note] The table shows effect heterogeneity in labor markets with a single TV station. Sources and specification as in baseline.

Part I

ONLINE APPENDIX

9 APPENDIX: Extensions

9.1 Equilibrium of the Superstar Model

Each stage manager maximizes profits by hiring a worker with talent t_p, taking its own firm characteristic as given. It will be convenient to express the hiring decision as choosing a percentile p from the talent distribution. Hence, in the optimization of manager i we can write the production function $Y(S,t)$ as $Y_i(p)$. The firm problem is therefore given by:

$$max_p Y_i(p) - w(p)$$

where $w(p)$ is the wage for a worker at percentile p of the talent distribution. Also, I extend the model and allow for entry and exit. This gives rise to a fourth equilibrium object, the participation threshold $\bar{p}$.

Condition 1 is a consequence of the single crossing condition $Y_{st} > 0$.[40] To derive condition 2 I start from the fact that the equilibrium is incentive compatible. Incentive compatibility guarantees that for each firm i the optimal worker p meets:

$$Y_i(p) - w(p) \geq Y_i(p') - w(p') \quad \forall \ p' \epsilon [0,1] \tag{9}$$

The number of IC constraints can be reduced substantially. If the IC holds for the adjacent p' all the other ICs will hold as well. We can therefore focus on the percentiles just above and below p. The IC for the adjacent $p' = p + \epsilon$ can be further simplified if Y is differentiable in p. Divide equation 9 by ϵ and let $\epsilon \to 0$.

$$\frac{w(p) - w(p+\epsilon)}{\epsilon} \leq \frac{Y(S_i,p) - Y(S_i,p+\epsilon)}{\epsilon}$$

$$w'(p) = Y_p(S_i,p) \tag{10}$$

The IC condition can thus be written as a condition on the slope of the wage schedule and proves result 2.

Participation constraints (PC) define the participation threshold $\bar{p}$. They guarantee that both firms and workers are staying in the industry. Denote the reservation wage of workers $w^{res}$and the reservation profits ψ^{res} and hence the PC condition is:

[40] For a proof see for example Sattinger, 1975.

$$Y_i(p) - w(p) \geq \psi^{res} \quad \forall \ p\epsilon[\bar{p}, 1] \tag{11}$$

$$w(p) \geq w^{res} \quad \forall \ p\epsilon[\bar{p}, 1] \tag{12}$$

The marginal participant is indifferent between participating and hence the PC binds with equality: $w(\bar{p}) = w^{res}$ and $Y_i(\bar{p}) - w(\bar{p}) = \psi^{res}$. Individuals with lower levels of skill will work in an outside market where pay is independent of talent and given by w^{res}.

Summing over all firms, we can derive the total revenue in the economy: $S(\pi) = \int^{\bar{p}} h'(t)Y(\sigma(t), t)dt$. In equilibrium revenues equal total expenditure, denoted by $D(\pi)$, which delivers result 3. Supply is increasing in π (since as $\frac{\partial \bar{p}}{\partial \pi} < 0$), hence there is a unique market clearing price $\hat{\pi}$ as long as demand is downward sloping $D'(\pi) < 0$.

9.2 Skill Biased Technical Change and Pay Dispersion

The skill biased technical change model features two groups of workers, high (H) and low (L) skilled workers. To give the model the best possible shot at fitting the data assume that workers can have different amounts of H and L, call the quantity of skill t. Assume that t is distributed with an inverse CDF $h_H(t)$ and $h_L(t)$ respectively. Within a skill group workers are perfect substitutes and the firm therefore cares only about the total units of H and L employed. Production is given by a CES function with A_i the productivity of skill group i:

$$Y(H, L) = \left[A_H(\sum t^H)^\theta + A_L(\sum t^L)^\theta\right]^{1/\theta}$$

Because workers are perfect substitutes the law of one price applies. There is a single market clearing price for a unit of low and high talent, call them π_H and π_L. The price of high talent is given by:

$$\pi_H = A_H \left[\frac{\sum t^H}{Y}\right]^{\theta - 1}$$

And the wage of a high skilled individual with quantity of skill t^H is given by:

$$w_{t^H} = \pi_H \cdot t^H$$

Call the inverse CDF of wages p^w_{SBTC} and the probability that a wage is above w_p is:

$$p^w_{SBTC} = Pr(w_{t^H} > w_p) = Pr(t^H > \frac{w_p}{\pi_H}) = h_H(\frac{w_p}{\pi_H})$$

The top tail of the wage distribution follows the same distribution as t^H.[41] With Pareto shape parameter $\tilde{\alpha}$ the talent distribution is $h_H(t) = t^{-1/\tilde{\alpha}}$ and substituting this into the wage distribution yields a wage equation equivalent to the superstar distribution in 4 for the right value of $\tilde{\alpha}$.

9.3 Technical Change and Superstar Effects

This section derives proposition 2.1.

Part a) to see how growth varies across the distribution, differentiate equation 5 wrt to the percentile of the distribution

$$\frac{\partial g_e}{\partial p} = \frac{\tilde{\lambda}\tilde{\xi}}{\lambda\xi}\frac{(\kappa-1)}{\kappa}\frac{\xi}{\alpha}w^{\frac{(\kappa-1)}{\kappa}\frac{\xi}{\alpha}-1}\frac{\partial w}{\partial p} > 0$$

Since wages are increasing along the wage distribution, the expression on the RHS is positive. Hence, the growth rate increase as we move up the distribution, which proves part a). Note that an equivalent result holds for wage growth at the top.

Part b)The top income share is defined as the sum of incomes of individuals above percentile p divided by total income (G):

$$s_p = \int_p^1 w_j dj/G$$

We showed above that wages follow a Pareto distribution with shape parameter $\lambda = \frac{\alpha}{\xi}$. Note that for a Pareto distributions the top income share is given by $s_p = (1-p)^{1-\lambda}$, with λ^{-1} the shape parameter of the distribution.[42] The growth in the top income share from superstar effects is therefore given by:

$$g^{s_p} = \frac{s_p^{t+1}}{s_p^t} \approx \frac{(1-p)^{1-\frac{\xi}{\kappa\alpha}}}{(1-p)^{1-\frac{\xi}{\alpha}}} = (1-p)^{-(\kappa-1)\frac{\xi}{\kappa\alpha}}$$

[41] Here we assume that low skill workers do not features in the top tail of the wage distribution.

[42] Even for variables that do not follow a Pareto distribution, there is still a lambda now varying with p. Many income variables are approximately Pareto and lambda is only slowly varying and the result holds approximately. This result has been used extensively to calculate top income and wealth shares.

The second step uses the property of a Pareto variable, the approximate result indicates that this results can works approximately for a much wider set of distributions. The final equality cancels terms. Top incomes shares are growing and we can see that the growth rate is increasing in p. This implies that the income share of the top 0.1% growths faster than the share that goes to the top 1%, which in turn growths faster than the share of the top 10%. The top 1% takes home a growing fraction of the income among the top 10%.

The core result, that a unit of talent becomes more valuable, holds independent of the distributional assumptions. As it becomes feasible to serve bigger markets, the wage-talent profile pivots and becomes steeper. For the general case we can show this by differentiating condition 2 with respect to s:

$$w_{ps}(t^*) = Y_{ps}(t^*) + Y_{pp}(t^*)\frac{\partial t}{\partial s} = \frac{w''(t^*)}{\theta'(t^*)} > 0 \tag{13}$$

The second equality uses positive assortative matching to invert the assignment function $t^* = \sigma^{-1}(s)$ and differentiates to yield $\frac{\partial t}{\partial s} = \frac{1}{\sigma'(t)}$. The effect of market size on the wage slope is positive. This follows from the convex wage schedule discussed above and the positive assortative matching of talent and market size. We don't need to appeal to the envelope theorem here. The envelope theorem doesn't apply in an assignment model. An employer who increases the market size is able to poach a better worker from a competitor and thus has first order effects on other market participants. Even without appealing to the envelope theorem we can sign the equation as long as the assignment function is invertible.

Part c) The falling wage is a result of the growing supply of talent, which reduces π. As a result the Pareto scale parameter in equation 4 falls ($\lambda' < \lambda$) and the wage distribution shifts inward.[43] This level shift occurs across the distribution, among stars the growth in returns from scalability over-compensates for the fall in π, but for non-stars the decline in π dominates. This effect is also reflected in the share of jobs with mid pay. Given the assumption on the demand elasticity $(1 - \varepsilon < (\gamma\phi\pi)^{\kappa-1})$, the first term of equation 5 is smaller than 1 (ie $\frac{\tilde{\lambda}\tilde{\xi}}{\lambda\xi} < 1$). As $w \to 1$, the growth rate $g_e = \frac{\tilde{\lambda}\tilde{\xi}}{\lambda\xi} - 1 < 0$ and hence the share of jobs at such pay levels is declining.

Part d) In the model with entry and exit the participation constraint (PC)

[43] If we maintain that the outside option is fixed at a level b, the lowest wages are fixed at b and adjustment occurs through exit rather than falling wages. Wages at the bottom could decline if there is a cost to exiting, for example search costs, or if payoffs from the outside option also fall.

ensures that the marginal participant is indifferent between working and the outside option: $w(\bar{p}) = w^{res}$. Deriving the equilibrium wage from integrating 2, we get $w(p) = \int_{\bar{p}}^{p} Y_t(\sigma(t), t) dh(t)$. And hence: $w(\bar{p}) = Y_t(\sigma(\bar{p}), \bar{p}) = w^{res}$. When Y_t fluctuates changes in $\bar{p}$ ensure that the PC holds. Raising $\bar{p}$ implies that the returns for the marginal worker $Y_t(\sigma(\bar{p}), \bar{p})$ increase, since $Y_{tt} + Y_{ts} > 0$. Hence when falling talent prices ($\tilde{\pi} < \pi$) lead to a decrease in Y_t, equilibrium requires that the participation threshold increases. Periods of technical change therefore lead to higher $\bar{p}$, which confirms statement d).

9.4 Technical Change and SBTC Models

9.4.1 Proportional Top Income Growth

Skill biased technical progress makes high skilled workers more productive ($\tilde{A_H} > A_H$). The wage per talent unit therefore becomes:

$$\tilde{\pi_H} = \tilde{A_H} \left[\frac{\sum \tilde{t^H}}{\tilde{Y}} \right]^{\theta - 1} > \pi_H$$

Next consider wages. The baseline case assumes that labor supply is inelastic, hence the talent distribution ($h^H(t)$) is unchanged. Allowing for a labor supply response complicates notation and generates little additional insight.[44] The wages at p are given by:

$$p^w_{SBTC} = (\tilde{w}_p / \pi')^{-\frac{1}{\alpha}} \tag{14}$$

We now can show that technical change leads to very limited change in the distribution of wages. The growth of wages is given by:

$$g^w_p = \frac{\tilde{w_p}}{w_p} = \frac{\tilde{\pi_H}}{\pi_H} = g^w$$

Wage growth is the same across all percentiles in the top tail. At the top of the distribution technical change leads to a level shift in the wage schedule.

[44] The higher wage induces entry of workers where $\tilde{w}_t$ growths above the outside option b. These are workers with low levels of t and as a result the distribution of talent changes at the bottom end. For ordinary talent distributions this has little effect on the top tail of $G_H^{-1}(p)$. The result that follow therefore carry through approximately at the the top of the distribution.

9.4.2 No Fractile Inequality

With a skill biased demand shock the growth in the top income share is given by:

$$g^{s_p} = \frac{s_p^{t+1}}{s_p^t} = \frac{G^t}{G^{t+1}} \frac{\pi_{t+1} \int_p^1 p^{-\alpha} dp}{\pi_t \int_p^1 p^{-\alpha} dp} = \frac{g^{\pi}}{g^G}$$

The second step uses the definition of top income shares and equation 14. The final step collects terms and cancels. Top income shares grow as long as the price for talent growths faster than GDP. Strikingly, the growth rate of the top income share at p is independent of p. All top income shares are growing at the same rate. The ratio of the income share that goes to the top 1% and 10% is therefore unaffected by SBD shocks.

9.4.3 No Cannibalisation in SBTC Models

This section proofs that technical progress rules out falling wages in the SBTC model. I study a flexible SBTC model with arbitrary many skill groups 1 ... n. The production function is given by:

$$F(\alpha_1(\theta)L_1, \alpha_2(\theta)L_2, ..., \alpha_n(\theta)L_n)$$

Where L_i is type of labor i and α_i the associated productivity and θ is the driver of technical change. We allow for exit and therefore impose that no worker type is indispensable in production:

$$\frac{\partial F}{\partial L_i} < \infty \qquad \forall L_i$$

Technical change may affect different parts of the distribution differently, in particular we allow for extreme bias technical change that predominantly helps star workers. We do not ex-ante rule out that changes in technology reduces productivity for some types of workers. However, we impose that the overall effect of technology is positive, hence we assume there is no technical regress in production:

$$\frac{\partial F}{\partial \theta} = \sum L_i \frac{\partial \alpha_i}{\partial \theta} \frac{\partial F}{\partial L_i} > 0 \tag{15}$$

We want to show that this implies that:

$$\frac{\partial \alpha_i}{\partial \theta} \geq 0 \qquad \forall i$$

We proceed by contradiction and assume this was not the case, hence $\frac{\partial \alpha_i}{\partial \theta} < 0$ for some i. To see that this violates restriction 15, assume that all $L_j = 0$ for all $j \neq i$ and $L_i > 0$ for i. This implies $\frac{\partial F}{\partial \theta} < 0$, violating the assumption that technical progress cannot lead to falling productivity.

9.4.4 Extending SBTC to match Superstar Effects

It is useful to think abour changes to the SBTC model that are needed to replicate properties of superstar effects. The relation between the the two models is particularly apparent if we take the Rosen, 1981 superstar model where $f(s,t) = s \cdot q(s,t)$, hence output of an actor with quality t and audience s depends on the audience size s and the quality of the output produced q. Market clearing ensures that demand equals supply at price π and if spending is inelastic at K, market clearing becomes $\int \pi \cdot f(s,t)h'(t)g'(s)dsdt = K$. To define production function properties I use the following notation $\dot{q} = log(q)$ and derivatives are indicated by subscriptes. Assume $\dot{q}_s < 0, \dot{q}_t > 0$ and $\dot{q}_{sq} > 0$. Take a production function that meets this requirement, for instance:

$$q_i = t_i e^{1-\delta_t s_i}$$

the quality of output q is a product of talent (t_i) of individual i and crowdedness of the show s; the more exclusive the show, the higher is the utility from it. Having Madonna play at a private dinner party brings greater utility, compared to listening to the same song on a recording. Comparative advantage ($\dot{q}_{sq} > 0$) implies that a world class performer is better able to deal with bigger audiences and thus quality suffers less from crowding $\partial\delta_t/\partial t < 0$. Take the simple case where $\delta_t = t^{-\phi}$, and ϕ is the technology that determines how easy it is to deliver a performance to a large audience. An increase in ϕ, as before, implies it gets easier to scale productions. Wages are given by $w(s,t) = P(q(s,t))s$, with P the price charged for a show of quality q. Profit maximizing implies equilibrium wages are given by:[45]

[45] with equilibrium $s = -q/q_s$. Notice that wages feature the "talent amplification effect," where wages are more dispersed than talent, as long as $\phi > 1$. The intercept term captures the market value of a talent unit and does not vary by talent, it therefore is a level shifter in wages.

$$\dot{w}_i = \dot{\pi} + \dot{t}_i\phi$$

Superstar effects arise when ϕ increases, hence when it gets easier to reach big audiences. Wages are affected by this in two ways:

$$\frac{\partial \dot{w}}{\partial \phi} = \dot{\pi}_\phi + \dot{t}_i$$

the first term (π_ϕ) is identical for all types of workers, while the later term (t_i) is bigger for more talented actors. The first term captures the cannibalisation effect that hurts all workers as $\pi_\phi < 0$, but this particularly affects the less talented actors who don't simultaneously benefit from the offsetting scale effects. As a result of the technical shift wage dispersion goes up, notice that this result relates to the dispersion in *log* wages. This log wage dispersion is an important difference to standard models of skill biased demand, where workers are paid in line with their productivity: $w_i = at_i$. A skill biased demand shock, which increase the skill premium ($a \uparrow$), changes the wage distribution, but has only a level effect on the dispersion in *log* wages. SBTC therefore do not generate dispersion in log wages. The SBTC model can replicate superstar effects if we make each skill type unique. Hence, rather than assuming that people have different levels of skill, we assume that each worker is a specific skill group. The crucial difference is that it makes all workers imperfect substitutes and thus allows wages across individuals to differ by more than there skill units. Simultaneously, this makes it feasible for wage gaps to grow differentially for workers with the same skill discrepancy and thus we can make a marginal talent unit more valuable at the top and generate superstar effects. The clue is that imperfect substitutability breaks the law of one price which forced wage differences to be proportional to skill differences. Each individual has it's own productivity term and the wage of i is given by:

$$w_i = A_i \left[\frac{1}{Y}\right]^{\theta - 1}$$

we can thus replicate superstar effects if a technical change generates $\frac{\partial \dot{A}}{\partial \phi} = \dot{\pi}_\phi + \dot{t}_i$. A model with a continuous distribution of unique talent types replicates, but does not coincide with the superstar model. A key difference is the process that generates wage dispersion. In the SBTC model wage inequality changes from biased productivity shifts, while the superstar model is more parsimonious and rising inquality is the result of the more specific SRTC.

There are two unappealing feature of extending the SBTC model in the way described above. First, the model becomes overly general – we have as many productivity terms as workers and thus can explain any kind of wage change, making it a somewhat uninteresting model. Second, to explain falling wages we require technical regress. This requires a counterintuitive scenario where lower skilled workers lose access to the previous, more productive technology. In other words innovation makes them "forget" how to be productive.

10 APPENDIX: Empirics

10.1 Summary Statistics

Table 9 reports summary statistics for the baseline local labor market sample. This covers the 722 local labor markets for 4 Census years and thus 2,888 observations. The first set of results report statistics on the availability of television. Television was unavailable in the first decade and becomes available in later decades. For filming however, the advent of national filming leads to a decline. Averaging over the full sample period there is only 0.02 TV stations filming in an average local labor market. This of course hides large variation across time and space capturing regional heterogeneity and the rise and disappearance of local TV filming discussed in the text. Data for show audience and revenues is only available for a subset of local labor markets due to missing data. The next set of statistics cover local entertainer labor markets. The average local labor market employs 177 performance entertainers. But there is considerable heterogeneity across local labor markets, stemming from the difference in the population of local labor markets (see demographics). Employment in all other leisure related activities – include drink & dine professions, as well as interactive leisure activities – is about 2,500 individuals in an average local labor market. The 99th percentile of the entertainer wage distribution average close to $5,700. As described in the text, this value is only computed for the larger local labor markets. Data on county fairs reports average attendance and spending in three categories: tickets, shows, rides & carnival. These data show that county fairs are a popular event, the average fair attracts about 25,000 visitors. This data is available at higher frequency and spans over 8,000 local labor market-year observations. Finally, the table reports demographic information on the population in the local labor markets. The average local labor market has 229,000 inhabitants and 86,000 workers, earning on average \$1,698. Median income is missing for one observation.

10.2 Robustness checks

10.2.1 Top Income Metrics

The baseline outcome variable normalizes the number of top earners by aggregate employment in entertainment. This has the convenient effect that the result is a percentage change. The numerator doesn't vary at the local labor market level, changes in this variable should therefore be captured by the year fixed effect. We may however worry that since the variable enters multiplicatively, the additive year fixed effect doesn't completely control for changes in the denominator. In column 2 Table 12 I therefore re-run the baseline regression using the count of top earners as outcome. In an average labor market 18 individuals are in the top percentile. TV more than doubles the number of top earners. Column 1 repeats the baseline regression. The normalization changes the units of the results, but the basic conclusion remains unchanged. This confirms that the normalization has no substantive effect on the result.

Figure 14 illustrated the evolution of various alternative top income measures. The figure shows the the 99th percentile of the Census wage distribution over time. This is the threshold that defines top earners in the baseline estimates. The figure contrasts this threshold with alternative top income thresholds. These include the thresholds calculated by (Piketty & Saez, 2003) and the 95th percentile of the wage distribution and the 95th percentile of the entertainer wage distribution. All of these are below the wage top-code applied in the data. The series move similarly. In practice it will therefore matter little how a top earner is defined. Table 12 confirms this formally. It repeats the previous analysis using other top income measures. Column 1 repeats the baseline estimate. Column 3 uses the top income percentile as defined by (Piketty & Saez, 2003). With this definition of top earners slightly more entertainers are top earners. The effect of TV remains however unchanged. The the number of people in the top percentile about doubles.

Column 4 and 5 look at the wage distribution among entertainers. By definition 1% of entertainers will earn wages above the 99th percentile of the entertainer wage distribution. Mechanically the share of top earners thus can't change. Instead the analyses looks at where these individuals live. If TV had a positive effect on top incomes, the number of top earning entertainers increases in areas where TV productions are filmed and declines elsewhere. With the Census data it is not possible to analyze the 99th percentile of the entertainer wage distribution This value is above the top code in some years. While we saw that the 99th percentile of the overall wage distribution stays below the top code, the same doesn't hold true in entertainment wage distribution because entertainer wages are more skewed than overall wages. The analysis therefore looks at entertainers above the 95th percentile of the entertainer wage distribution. Analyzing within entertainer wage dispersion has the appealing advantage that it is a measure of inequality in the affected sector. This measure is however problematic if TV induces substantial exit in the entertainment sector. Exits would shift the 95th percentile even in the absence of any effect of television on top earners. If television results in an exit of the bottom 10% of entertainers, the 95th wage percentile would rise. If there was no further effect on top earners, we would find that fewer entertainers are top earners after the introduction of television. Hence, this measure will lead to a downward biased in the estimate of TV. Indeed in column 3 the number of top earners increases by less. The increase here is 20% over the baseline. To address the endogeneity issue column 4 keeps the 95th percentile fixed at the 1940 level. This measure is thus unaffected by exit of entertainers. This estimate is indeed substantially bigger than column 3. These results confirm that television led to a substantial increase in top earnings in entertainment.

10.2.2 Pre-Trend

A challenge for estimating pre-trends with this sample is that wage data in the Census is first collected in 1939. Since the Census is decennial this only allows for a single pre-treatment period. To estimate pre-trends I therefore combine the Census data with data from Internal Revenue Services (IRS) tax return data. In 1916 the IRS published aggregate information on top earners by occupation-state bins. Data for actors and athletes are reported. I link the Census data with the tax data and run the regressions at the state level. Table 14 reports the results. Column 1 repeats the baseline estimate with data aggregated at the state level. Despite the aggregation at the state level the effect remains highly significant. Column 2 adds the additional 1916 data from the IRS. The results stay unchanged. Column 3 shows the differences in top earners in treatment and control group for the various years. It shows a marked jump up in top earners in the treated group in the year of local TV production. The coefficient on the pre-trend is not significant because the standard errors are large. If anything the pre-period saw to a decrease relative decrease in top earners in the treatment areas. Even if taken at face value the pre-trends thus can't explain the identified positive effect of TV.

10.2.3 Placebo Occupations

Television only changed the production function of a handful of occupations, we can therefore use alternative occupations as placebo group. The ideal placebo group will pick up changes in top income in the local economy. The main high pay occupations are therefore used as placebo group, these professions are medics, engineers, managers and service professionals. If TV assignment is indeed orthogonal to local labor market conditions, we would expect that such placebo occupations are unaffected. Results for the placebo group are reported in 15. TV does not show up in top pay of the placebo occupations. The only occupation group with a significant positive effect are performance entertainers. Column 1 shows that the placebo group doesn't experience any growth in top incomes. Moreover, the estimated effect on performance entertainers remains similar to the baseline in Table 15. Column 2 allows for separate impact of television across the different placebo occupations. Only performance entertainers experience the significant and large top earner rise.

With the inclusion of the placebo occupations, I can run a full triple difference regression. In this specification there are treated and untreated workers within each labor market. We already controlled for location specific trends before, this specification will go further and allow for a non-parametric location specific time fixed effect. An example where this might be necessary is if improved local credit conditions result in greater demand for premium entertainment and simultaneously lead to the launch of a new TV channel. This may lead to an upward bias in the estimates. My treatment now varies at the time, labor market and occupation level. This allows me to control for pairwise interactions of time, market and occupation fixed effects. These will address the outlined credit access problem as the fixed effects will now absorb location specific time effects.

Column 3 shows the results. The effect on performance entertainers remains close to the baseline estimate. The additional location specific time and occupation fixed effects therefore don't seem to change the findings. This rules out a large number of potential confounder. The introduction of a "superstar technology" thus has a large causal effect on top incomes and this effect is unique to the treated group.

10.2.4 Quantile Regressions

A further method of testing the effect of TV across the distribution is through quantile regressions. A number of recent papers have extended the use of conditional quantile regressions to panel settings. In the linear regression framework additive fixed effects lead to a "within" transformation of the data. In the non-linear quantile framework additive linear fixed effects will not result in the standard "within" interpretation of the estimates. Adding fixed effects may therefore not be sufficient for identification. Chetverikov, Larsen, and Palmer (2016) develop an quantile estimator that handles group level unobserved effects if treatment varies at the group level. Similarly, Powell, 2016 develops a panel quantile estimator that mimics the "within" transformation of fixed effects for the quantile regression.

A shortcoming of the quantile regression is that the estimates are sensitive to entry and exit. The magnitude of the quantile effect is therefore hard to interpret. However, the relative magnitude across percentiles is still informative and the test relies exclusively on such relative patterns. Recall that SBD predicts a homogeneous growth rate, while the superstar model predicts larger wage growth rates at the top. To test whether either model matches the data, I run quantile regressions at various percentiles. I restrict myself to quantiles for the median and above since the results were derived by using an approximation for the top of the distribution. I follow the procedure in Chetverikov et al. (2016) to implement the difference in difference for quantile regressions. The estimated coefficients are plotted in figure 16, alongside the prediction of the SBD model. The effect is biggest at the top of the distribution and effects are notably smaller at the lower percentiles. This result is in line with the superstar model but contradicts a model of SBD. Table 16 reports the panel quantile estimates using the Powell, 2016 approach.

10.3 Data construction

10.3.1 Local labor markets

The analysis defines local labor markets as commuting zones (CZ). A labor market is an urban center and the surrounding commuters belt. The CZs fully cover the mainland US. The regions are delineated by minimizing flows across boundaries and maximizing flows within labor markets, they are therefore constructed to yield strong within-labor-market commuting and weak across-labor-market commuting. David Dorn provides crosswalks of Census geographic identifiers to commuting zones (Autor

& Dorn, 2013). I use these crosswalks for the 1950 and 1970 data

I build additional crosswalks for the remaining years. For each Census I use historic maps for the smallest available location breakdown. I map the publicly available Census location identifiers into a commuting zone. No crosswalk is available for the 1960 geographic Census identifier in the 5% sample and the 1940 Census data. Recent data restoration allows for more detailed location identification than was previously possible (mini-PUMAs). To crosswalk the 1940 data, I use maps that define boundaries of the identified areas. In GIS software I compute the overlap of 1940 counties and 1990 CZ. In most cases counties fall into a single CZ. A handful of counties are split between CZ. For cases where more than 3 percent of the area falls into another CZ, I construct a weight that assigns an observation to both commuting zones. The two observations are given weights so that they together count as a single observation. The weight is the share of the county's area falling into the CZ. The same procedure is followed for 1960 mini PUMAs. Carson city county (ICSPR 650510) poses a problem. This county only emerges as a merger of Ormsby and Carson City in 1969, but observations in IPUMS are already assigned to this county in 1940. I assign them to Ormsby county (650250). CZ 28602 has no employed individual in the complete count data in 1940.

10.3.2 Worker data

Data is provided by the Integrated Public Use Microdata Files (IPUMS, Ruggles, Genadek, Goeken, Grover, and Sobek, 2017) of the US decennial census from 1930-1970 (excluding Hawaii and Alaska). I use the largest publicly available sample for each Census, for 1970 I combine form 1 and form 2 metro samples to obtain the most granular spatial data. Extending the time period in either direction is precluded by changes in variable definitions. Prior to 1930, the Census used a significantly different definition of employed workers than in my period of interest, and from 1980 onwards, the Census uses different occupation groups. Most variables remain unchanged throughout the sample period. IPUMS has taken great care to provide consistent measures of variables that did change.

- there are 722 commuting zones (CZ) covering the mainland USA. These regions are consistently defined over time.
- there are 28 relevant occupations. 1950 occupation codes are

  - Treatment group: 1, 5, 31, 51, 57
  - Placebo group: 0, 32, 41, 42, 43, 44, 45, 46, 47, 48, 49, 55, 73, 75, 82, 200, 201, 204, 205, 230, 280, 290, 480

- Aggregates are calculated using the provided sample weights
- variables used incwage, occ1950 (in combination with empstat), wkswork2, hrswork2
- To match TV signal exposure to the Census I map county level TV signal information onto geographic units available in the Census. The geographic match uses the boundary shapefiles provided by NHGIS (Manson, Schroeder, Riper, and Ruggles, 2017). I then identify how many TV-owning households are in each TV station's catchment area. This allows me to construct a measure of potential audience size.

10.3.3 Employment

Number of workers are based on labforce and empstat. Both variables are consistently available for 16+ year olds. Hence the sample is restricted to that age group. Occupation is recorded for age>14. I use this information for all employed. This is available consistently with the exception of institutional inmates who are excluded until 1960. The magnitude of this change is small and the time fixed effect will absorb the effect on the overall level of employment. The definition of employment changes after the 1930 Census. Before the change, the data doesn't distinguish between employment and unemployment. In the baseline analysis I therefore focus on the period from 1940 onwards. For this period the change doesn't pose a problem. An alternative approach is to build a harmonized variable for a longer period, this includes the unemployed in the employment count for all years. I build this alternative variable and perform robustness checks with it. The results remain similar. For two reasons the impact of this change on the results is smaller than one might first think. First, most unemployed don't report an occupation and thus don't fall into the sample of interest.[46] Second, the rate of unemployed is modest compared to

[46] There are a number of cases were the unemployed report an occupation. This occurs if they have previously worked. I construct an employment series that includes such workers for the entire sample period. This measure is a noisy version of employment as some job losers continue to count as employed. Since the share of these workers is small, the correction has only small effects on the results.

employment and thus including them doesn't dramatically change the numbers.

I use the IPUMS 1950 occuption calssification (Occ1950). This data is available for years 1940-1970. For previous years the data is constructed using IPUMS methodology from the original occupation classification. Occupational definitions change over time. IPUMS provides a detailed methodology to achieve close matches across various vintages of the US census. Luckily the occupations used in this analysis are little affected by changes over time. More details on the changes and how they have been dealt with are: The pre 1950 samples use an occupation system that IPUMS judges to be almost equivalent. For those samples IPUMS states: "the 1940 was very similar to 1950, incorporating these two years into OCC1950 required very little judgment on our part. With the exception of a small number of cases in the 1910 data, the pre1940 samples already contained OCC1950, as described above." For the majority of years no adjustment all is therefore necessary. Changes for the 1950-1960 period - Actors (1950 employment count in terms of 1950 code: 14,921 and in terms of 1960 code: 14,721), other entertainment professions are unaffected. Changes from 1960-1970: Pre 1970 teachers in music and dancing were paired with musicians and dancers. In 1970 teachers become a separate category. My analysis excludes teachers and thus is unaffected by this change. Athletes disappear in 1970 coding. The analysis therefore only uses the athlete occupation until 1960. The only change that has a major effect on worker counts is for "Entertainers nec". In 1970 ca. 9,000 workers that were previously categorized as "professional technical and kindred workers" are added and a few workers from other categories. The added workers account for ca. 40 percent of the new occupation group. The occupation specific year effect ought to absorb this change. I have also performed the analysis excluding 1970 and find similar results. Moreover I find the TV effects for each occupation individually. The classification changes therefore seem to have little effect on the results.

The industry classification also changes over time. I use the industry variable to eliminate teachers from the occupations "Musicians and music teacher" and "Dancers and dance teachers." The census documentation does not note any change to the definition of education services over the sample period, however the scope of the variable fluctuates substantially over time. From 1930 to 1940 the employment falls from around 70,000 to 20,000, from 1950 to 1960 it increases to around 200,000 and falls back to around 90,000 from 1960 to 1970.

The control group are workers in top earning professions outside entertainment (lawyer, medics, engineers, managers, financial service). The relevant occupations are available across most years. Exceptions are 1940 where a few occupations in engineering, medicine and interactive leisure are grouped together and in 1970 where the floor men category is discontinued. I control for those changes with year-occupation fixed effects in the regressions. The effects occur within occupations rather than between them, results for all occupations separately are available upon request.

10.3.4 Wage data

Labor earnings are used to be consistent with the model (wages, salaries, commissions, cash bonuses, tips, and other money income received from an employer). This differs from Piketty et al who use earnings data of tax units. As described above, I use wage data and focus on individual data rather than earnings of a tax unit. This choice makes economically sense for this setting. The superstar theory is concerned with individual labor earnings and abstracts from household composition and capital income.

The data is from the US Census with the following sample features. Census data on wages refer to the previous calendar year. In 1940 and 1960+ every individual replies to this question - in 1950 only sample line individuals do (sub-sample). The 1940 100% sample is not top coded, other years are. The 99th percentile threshold is always below the top code, hence the top code doesn't pose a problem here. I calculate measures for top income dispersion in entertainment for each market by year. Some measures, for instance income dispersion, are not additive across occupations and I therefore calculate a single dispersion coefficient per year-local entertainer labor market, which pools the micro data for the five occupations affected by TV. Wage data is in real 1950 terms

10.3.5 Pareto Interpolation

Top income shares can be computed straight from the data if the full population is covered. Without information on the full population the standard approach in the literature is to use Pareto approximations (e.g. Atkinson and Piketty, 2010; Atkinson, Piketty, and Saez, 2011; Blanchet, Fournier, and Piketty, 2017; Feenberg and Poterba, 1993; Kuznets and Jenks, 1953; Piketty and Saez, 2003). This assumes that the income distribution is locally Pareto and interpolates incomes between two observed individuals, moreover it allows to extrapolate the top tail of the distribution. In a

Pareto distribution two parameters, pin down the wage distribution. In practice there are a number of challenges. Key to the dispersion is the "Pareto coefficient." There are at least four challenges in estimating the parameter. The first is misspecification, we do not belief that wages exactly follow a Pareto distribution. Second, outcomes are an order statistic which violates the iid assumption. Third measurement error in wages affects the regressor. Fourth in samples the population rank of an observation is not observed. I address these issues by analyzing the performance of popular methods in years where the full population data allows for validation.

The beauty of the Pareto distribution is that a it is a straight line in the log space. This holds because the CDF of a Pareto distribution is linear in logs: $1 - F(w) = (w/\omega)^{-1/\alpha}$. Once we know two points on the line we can reconstruct the slope and intercept of the line and have fully characterized the distribution. The slope captures the "Pareto coefficient". The slope is given by: $\alpha_{i,j} = [ln(income_i) - ln(income_j)] \,/\, [ln(rank_i) - ln(rank_j)]$. Since we usually observe many points we could calculate many Pareto coefficients and combine them in an optimal way. Fortunately economist have thought about the best way of fitting a line through a cloud of points. We can fit a line to estimate the Pareto coefficient by running a regression of the form[47]:

$$ln(income_i) = \beta - \alpha \cdot ln(rank_i) + \epsilon_i$$

It turns out that OLS is a poor approach here. The Gauß Markov assumptions are violated making OLS inefficient and bias. The outcome variables are order statistics, resulting in heteroskedacticity and correlation of errors across observations. Moreover, the log transformation implies that $E(\epsilon_i) = E(log\varepsilon_i) \neq 0$, making OLS biased. The latter problem can be addressed by replacing the regressor with the Harmonic index (Blanchet, 2016). And efficiency can be achieved with MLE.[48] Polivka, 2001 and Armour, Burkhauser, and Larrimore, 2015 give an overview how MLE can be applied to this problem. A further challenge is misspecification. The Pareto distribution is used as an approximation and may not fit the data perfectly. In particular the distribution may fit better at the top than the bottom of the distribution. Even at the top of the distribution changing Pareto coefficients may be required to fit the data (Blanchet et al., 2017). Misspecification is particularly problematic for the

[47]Here $\beta = ln(\underline{income}) - ln(\underline{rank})$ where lower bars represent the lower bound of the interval considered

[48]Since the covariance structure of order statistics is known, GLS yields the same result

more efficient estimators (Finkelstein, Tucker, and Alan Veeh, 2006). I will test the performance of three estimators using real-world data by drawing samples from the full-count Census. This allows us to assess how estimators cope in data with i) small samples, ii) top coding and iii) bunching at tax thresholds and round numbers. I test the following estimators:

- Estimator with n total observations, T top coded observations, $rank_j$ the rank in the wage distribution (1 being the top), w_j wage at rank j and ω the smallest wage in the sample:
- MLE: $\hat{\beta}^{MLE} = \frac{1}{n}\sum_{j=1}^{n} log(w_j/\omega)$
- MLE (top code adjusted): $\hat{\beta}^{MLE_{TC}} = \frac{T}{n}\sum_{j=T}^{n} log(w_j/\omega) + T * log(w^{TC}/\omega)$
- OLS: $log(w_j) = \delta - \beta^{OLS} * ln(\frac{rank_j}{n+1}) + \epsilon_j$
- Close to cut-off: $\hat{\beta^A} = (\sum_{j=1}^{3} \frac{ln(w_j/w_{j-1})}{ln(rank_j/rank_{j-1})})^{-1}$
- Extrapolation: The standard method of calculating top income shares fits a Pareto curve through the observed data and computes income shares as area under the curve. For the Pareto distribution the fraction that falls in the tail is captured by a single Parameter. We can thus compute any top income share once we know the tail index of the Pareto distribution. For other distributions the tail index varies for different percentiles, in that case we have one shape parameter that allows to compute the top 1% income share and a different one to compute the top 0.1% share. A well known feature of extreme value theory is that in the the tail many regular distribution only differ by a slow moving function from the Pareto. Using the Pareto parameter estimate just below the cut-off may thus yield a reasonable approximation even if the data generating process is not Pareto.

Table 17 shows the results. They suggest that OLS and MLE perform relatively poorly in small samples of the data of interest. I find that the best performing estimator is the average of the alpha values just below the top code. The difference to OLS and MLE estimates is the weight attached to values far from the top-code. OLS and MLE give a non zero weight to observations further away from the top-code. This approach will yield greater bias if the Pareto distribution is not a perfect fit and observations far from the top-code are poor proxies for the distribution beyond the top-code. Consistent with this, I find that the OLS and MLE perform worse in smaller samples. For the application here I therefore focus on Pareto interpolation based on observations closest to the top-code. It should be stressed that this result is specific to the data in this context. More general results for Pareto inference with real-world data should be conducted to establish the wider relevance.

For each local labor market and year I derive the Pareto coefficient. At the bottom of the income distribution the Pareto distribution has been found be a poor fit, I therefore discard Pareto parameters based on observations at the bottom quarter of the distribution. The results are however robust to including those observations. Next, I use the local labor market- year specific Pareto coefficient to estimate top income shares. Here I make use of the fact that for a Pareto distribution top income shares are given by: $S_{p\%} = (1-p)^{\frac{\alpha-1}{\alpha}}$.

10.3.6 Data on Market Reach of Entertainment Shows

Data on potential show audiences is collected from the "Julius Cahn-Gus Hill theatrical guide." For each local labor market I compute the potential maximum audience. For physical venues this is the seating capacity of the largest venue.

Show revenues in theatres are the price of tickets multiplied by the audience. I use the average price if multiple ticket prices are reported. For TV shows I collect price data from rate cards. Such cards specify the price for sponsorship of a show at a local station, which allows me to compute the price charged for a TV show. From the price per show I can compute a price per TV viewer, analogue to a ticket price, which quantifies the marginal return to reaching one more customer. Price data is only available for a subset of observations. I infer prices based on a data from TV station ad-pricing in 1956 and theater ticket prices in 1919. I use them to estimate a demand elasticity for TV audiences, taking the supply of TV hours as given. The demand curve for a TV viewer is estimated as: $ln(price) = 4.051 + -0.460 * ln(TV households)$. The negative elasticity indicates that, as expected, the marginal value of reaching a household is declining. The negative demand elasticity in turn implies that TV station revenues do not increase 1:1 with audience, the revenue elasticity is 0.54.

The potential audience of TV shows is the number of TV households that can watch a local TV station. This is computed using information on TV signal catchment areas (from Fenton and Koenig, 2018) and TV ownership records from the Census.

10.3.7 Controls

Control variables are: share non white, male, high skilled (high school and above for people over 25) and median age and wage. Most variables are available consistently throughout the sample period. Income and education are only available from 1940 onwards. The race variable as has changing categories and varying treatment of mixed race individuals. I use the IPUMS harmonized race variable that corrects for those fluctuations were possible.

10.3.8 IRS Taxable Income Tables

Data from the Internal Revenue Service (IRS) allows me to extend income data backward beyond what is feasible with the Census.[49] To obtain records for entertainers, I digitize a set of taxable income tables that lists income brackets by state and occupation. The breakdown of the data by occupation and state is only available for the year 1916.

[49] Such tax tables have been used by Kuznets and Piketty to construct time series of top income shares for the US population.

References

Armour, P., Burkhauser, R., & Larrimore, J. (2015). Using the Pareto Distribution to Improve Estimates of Topcoded Earnings. *Economic Inquiry*.

Atkinson, A. B., & Piketty, T. (2010). *Top Incomes - A Global Perspective*.

Atkinson, A. B., Piketty, T., & Saez, E. (2011). Top Incomes in the Long Run of History. *Journal of Economic Literature*, *49*(1), 3–71.

Autor, D. H., & Dorn, D. (2013). The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market. *American Economic Review*, *103*(5), 1553–97.

Blanchet, T. (2016). Wealth inequality in Europe and in the United States: estimations from surveys, national accounts and wealth rankings. (June).

Blanchet, T., Fournier, J., & Piketty, T. (2017). Generalized Pareto Curves: Theory and Applications. *WID.world Working Paper*, *2017/3*(March).

Chetverikov, D., Larsen, B., & Palmer, C. (2016). IV Quantile Regression for Group-Level Treatments, With an Application to the Distributional Effects of Trade. *Econometrica*, *84*(2), 809–833.

Feenberg, D. R., & Poterba, J. M. (1993). *Income Inequality and the Incomes of Very High-Income Taxpayers: Evidence from Tax Returns*.

Fenton, G., & Koenig, F. (2018). The Labor Supply Response to Innovation in Entertainment: Evidence from TV. *Working paper*.

Finkelstein, M., Tucker, H. G., & Alan Veeh, J. (2006). Pareto Tail Index Estimation Revisited. *North American Actuarial Journal*, *10*(1), 1–10.

Kuznets, S., & Jenks, E. (1953). *Shares of Upper Income Groups in Savings*.

Manson, S., Schroeder, J., Riper, D. V., & Ruggles, S. (2017). *IPUMS NHGIS: Version 12.0*. University of Minnesota. Minneapolis.

Piketty, T., & Saez, E. (2003). Income Inequality in the United States, 1913-1998. *The Quarterly Journal of Economics*, *68*(February), 1–39.

Polivka, A. E. (2001). *Using Earnings Data from the Monthly Current Population Survey*.

Powell, D. (2016). Quantile Regression with Nonadditive Fixed Effects. *RAND Working Paper*, (April), 1–28.

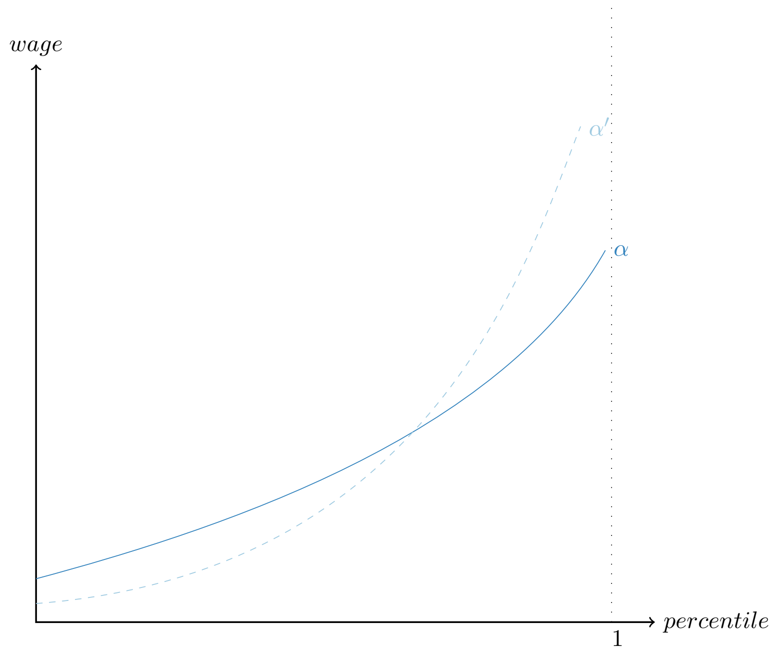
Rosen, S. (1981). The Economics of Superstars. *American Economic Review*, *71*(5), 845–858.

Ruggles, S., Genadek, K., Goeken, R., Grover, J., & Sobek, M. (2017). *Integrated Public Use Microdata Series: Version 7.0*. University of Minnesota. Minneapolis.

Sattinger, M. (1975). Comparative advantage and the distributions of earnings and abilities. *Econometrica*, *43*(3), 455–468.

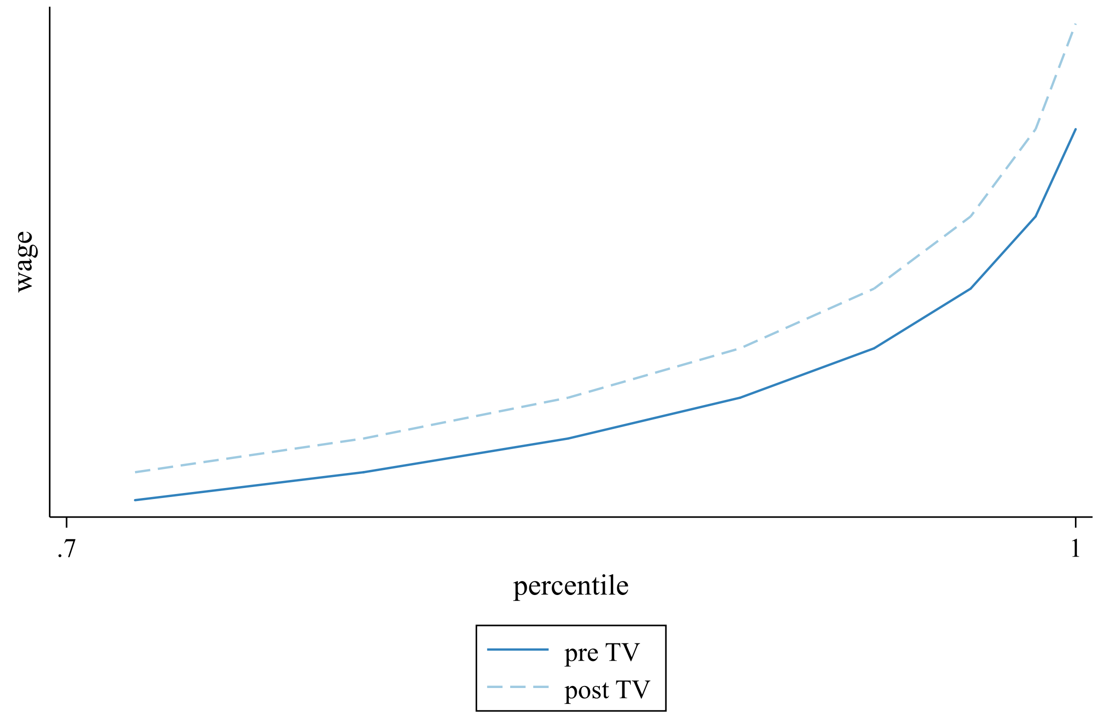
10.4 APPENDIX: TABLES & FIGURES

Figure 9: Superstar Wage Distribution



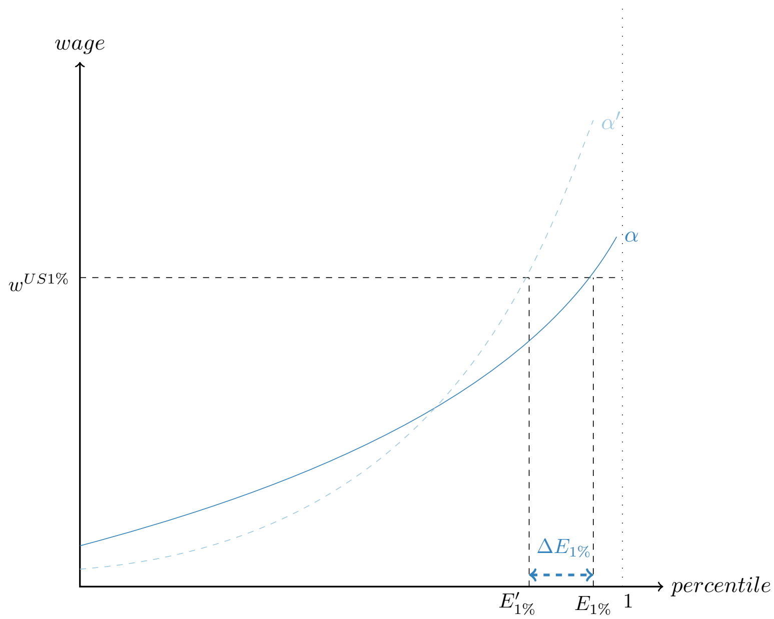
Note: Wages based on a superstar model ($w_p = \pi \cdot \kappa \cdot (1 - p)^{-(\alpha\gamma - \beta)}$). α is the shape parameter of the market size distribution ($\alpha' > \alpha$). The percentiles shown are the upper tail of the wage distribution. With exit they correspond to the percentiles in the pre-distribution.

Figure 10: Effect of Technical Change on Wage Distribution - Skill Biased Demand Model



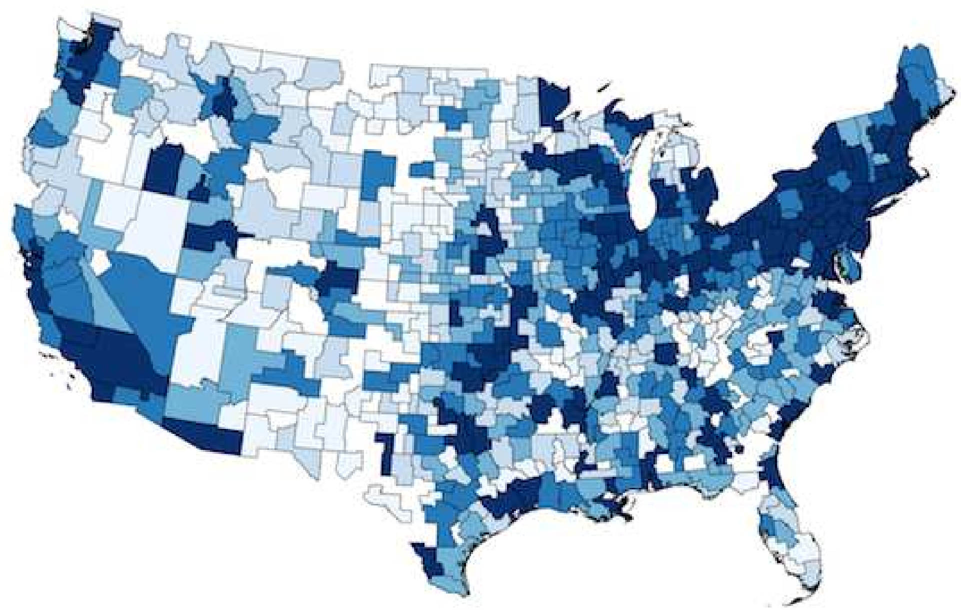
[Note] The figure shows the wage distribution above the 70th percentile. The talent distribution has been chosen to match the 1940 wage distribution. The change in the skill premium matches the growth in the share of top earners.

Figure 11: The Impact of Superstar Effects on Top Earner



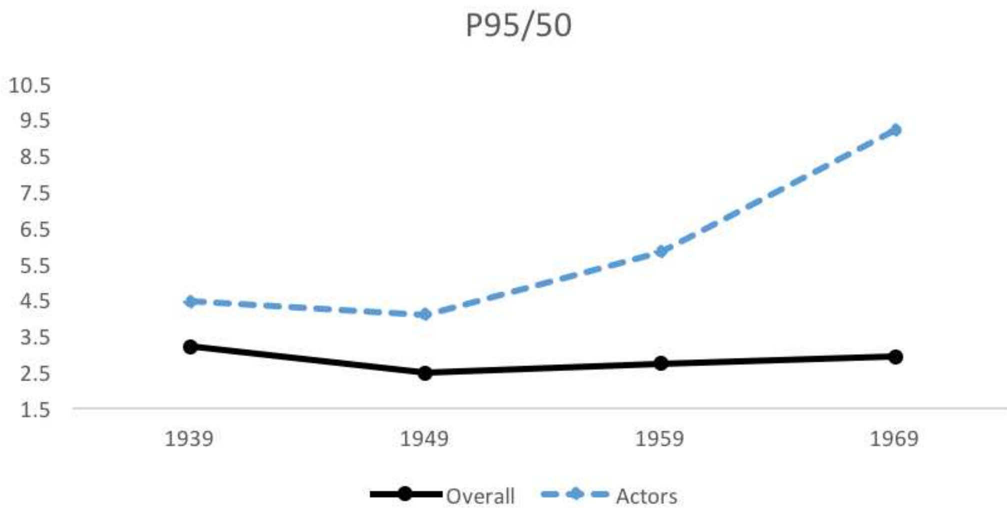
[Note] Details as in figure 9. $w^{US1\%}$ is a wage threshold that defines a top earner, e.g. the national top percentile. $E_{1\%}$ and $E'_{1\%}$ are the share of entertainers above the threshold. $\Delta E_{1\%}$ is the change in top earners when market size becomes more dispersed (move from α to α').

Figure 12: Theatre Seating Capacity



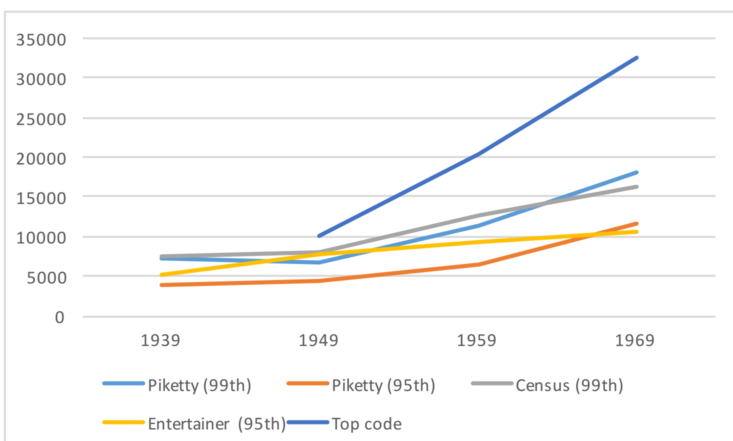
[Note] Performance venues are the venues listed in Julius Cahn-Gus Hill's 1921 theatrical guide. Size refers to the average seating capacity of the largest venues in the commuting zone.

Figure 13: P95-P50 Gap



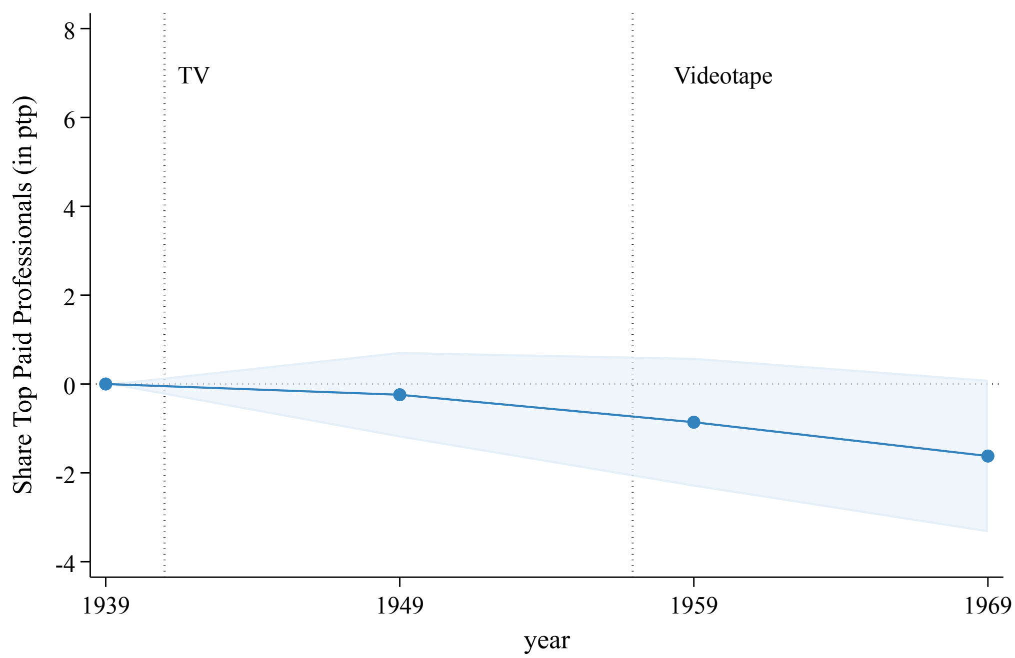
[Note] Figure reports the ratio of wages at the 95th and median. Percentiles are from the wage distribution reported in the US decennial Census for the lower 48 states.

Figure 14: Top Income Percentile Values



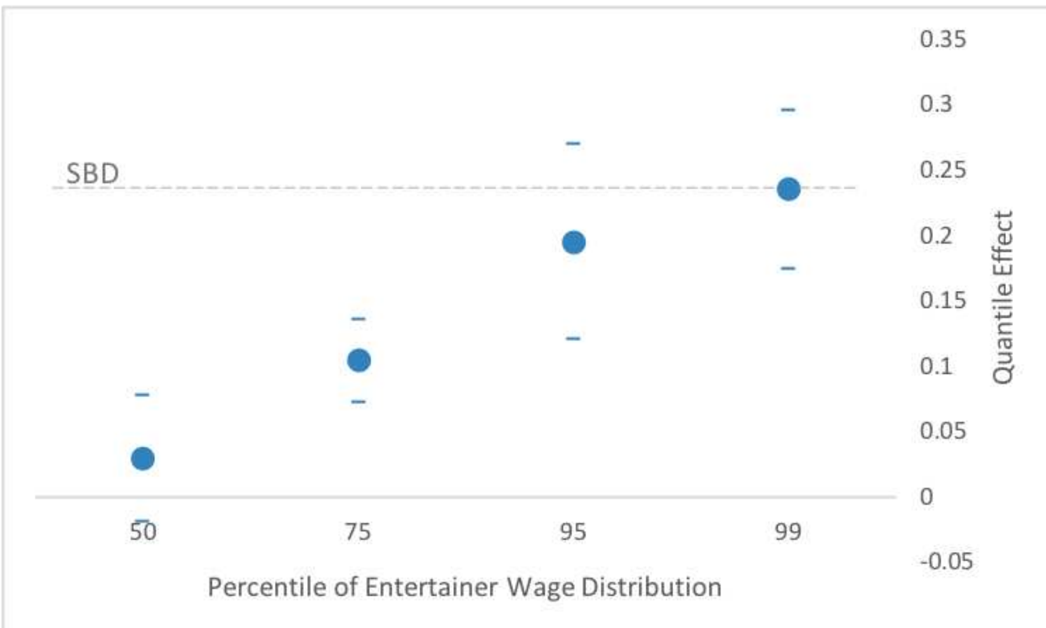
[Note] The Figure shows the top code cut-off in the US Census data and top percentiles of the wage distribution in the Census years. The name in the legend refers to the sourcee of the wage distribution: Census refers to percentiles in the Census data wage distribution, Entertainer to percentile in the distribution of entertainer wages in the Census, Piketty to the data reported in the World Top Income Database, top code is the top code in the IPUMS Census data – there is no top code for the 1939 full count Census data. The number in the bracket in the legend indicatees the percentile of the distribution that is shown.

Figure 15: Dynamic Treatment Effect of TV stations - Placebo Occupations



[Note] The figure shows regression coefficients from the dynamic difference in difference regression for placebo occupations. Reported are the coefficients on local TV antennas and 95% confidence bands are shown. Standard errors are clustered at the local labor market level.

Figure 16: Quantile Effects of Television



[Note] Each dot is based on separate quantile regression. The quantile regressions control for local labor market and year fixed effect. I use the technique developed in Chetverikov, Larsen, and Palmer, 2016 to do so. This amounts to calculating percentiles for each year-labor market observation and regressing those percentiles on the treatment. The first step uses the provided sample weights, while the second weights by cell size. If the top code bites for the analyzed percentiles, the cell is discarded. The dashed line represents the benchmark prediction of a skill biased demand model.

Table 9: Summary Statistics

	observation	mean	s.d.
		Television	
Local TV Stations	2,888	0.02	0.25
Local Filming Cost	2,888	0.14	1.36
Show Audience (thsd.)	2,656	72,811	66,719
Show Revenue (thsd. $)	2,656	4,182,516	3,834,174
TV Signal (%)	2,888	60	0.49
		Entertainment	
Employment in Leisure Activities	2,888	2,468	8,540
Employment in Performance Entertainment	2,888	177	936
Wage 99th Percentile of Entertainers ($)	1,435	5,704	4,576
Fair Visits (thsd.)	8,664	25	109
Fair Ticket Receipts (thsd. $)	8,664	2.94	1.89
Show Receipts (thsd. $)	8,664	1.64	0.97
Rides & Carnival Receipts (thsd. $)	8,664	0.92	7.50
		Demographics	
People (thsd.)	2,888	229	658
Worker (thsd.)	2,888	86	264
Median Income ($)	2,887	1,698	747
Population Density	2,888	2.5	7.8
Urban (%)	2,888	17	37
Minority (%)	2,888	9.6	13
Male (%)	2,888	50	2
Age	2,888	27.4	3.27

[Note] The table reports summary statistics for the 722 commuting zones over 4 decades. The data is decadal, except Fair data which is annual data 1946-1957. Show audience and revenue refers to the largest feasible shows in a CZ (see text for details), no data available for some CZs. Median income is missing in one CZ in 1940. Variables Urban Share and Filming Cost are held fixed throughout the sample. Source: US Census 1940-1970, Billboard magazine 1946-1956

Table 10: Effect of TV on Top Earner - Placebo Occupations

	(1)	(2)	(3)
	Panel A: ln(Wage at 99^{th} Percentile)		
Local TV station	0.023	0.019	0.016
	(0.004)	(0.003)	(0.005)
Outcome mean	9.08	9.08	9.08
Effect size	2.3%	1.9%	1.6%
	Panel B: Share of Occupation in US Top 1% (ptp)		
Local TV station	0.21	0.66	1.09
	(0.52)	(0.89)	(0.52)
Outcome mean	5.55	5.55	5.55
Effect size	4%	12%	20%
	Panel C: Local Population Share in US Top 1% (in 10,000)		
Local TV station	0.438	0.524	0.865
	(0.221)	(0.234)	(0.319)
Outcome mean	10.86	10.86	10.86
Effect size	4%	5%	8%
Cluster	722	722	722
Demographics	–	Yes	–
Local labor market trends	–	–	Yes

[Note] Each cell is the regression coefficient of a separate regression. Panel A uses a quantile regression for within group treatment Chetverikov, Larsen, and Palmer, 2016. For this procedure data is aggregated at the treatment level and uses 2,887 local labor market - year observations. Observations are weighted by cell-size, cells where 99th percentile cannot be computed are dropped. Panel B and C use a difference in difference regression and are based on respectively 62,042 and 62,746 observations at the occupation-local labor market - year level. The treatment is the number of TV stations in the local area. Reported baseline outcomes are the average of the dependent variable in treated areas in years without treatment. All regressions control for local labor market fixed effects, time fixed effects, local production cost of filming in years after 1956, in Panel B and C additionally for year - occupation fixed effects. The sample period spans 1940-1970. Demographics are median age, % female, % black, population density and trends for urban areas. The outcome variable in Panel B is the share of top paid entertainers calculated as described in the text, Panel C is the number of top paid entertainer devided by the population in a local labor market. Entertainer are Actors, Athletes, Dancers, Entertainers Not Elswhere Classified, Musicians. Observations are weighted by local labor market population. Standard errors are reported in brackets, they are clustered at the local labor market level.

Table 11: Effect of TV on Top Earner - Alternative Top Income Measures

	(1)	(2)	(3)
		Panel A: Count Entertainer in US top 1%	
Local TV station	30.91	32.09	19.31
	(8.92)	(9.92)	(8.31)
Outcome mean	15.53	15.53	15.53
		Panel B: Share Entertainer in US top 1% (denominator fixed)	
Local TV station	6.51	6.73	9.21
	(1.90)	(1.89)	(3.44)
Outcome mean	6.39	6.39	6.39
		Panel C: Percent US top 1% from Entertainment	
Local TV station	0.178	0.193	0.194
	(0.025)	(0.038)	(0.063)
Outcome mean	0.28	0.28	0.28
Cluster	722	722	722
Demographics	–	Yes	–
Local labor market trends	–	–	Yes

[Note] See table 1 Panel B denominator is the average number of entertainers per labor market in occupation o at time t. Denominator in Panel C is the total number of entertainers in local labor market c at time t.

Table 12: Alternative Top Income Measures

	(1)	(2)	(3)	(4)	(5)
	Share in US top 1%		Count top 1%	Share in top 5%	
Local TV station	90.19	132.5	30.91	31.64	120.0
	(26.25)	(35.92)	(8.92)	(16.36)	(47.85)
threshold	Census	Piketty & Saez	Census	Entertainer	Entertainer (1940)
mean outcome	94.27	109.09	18.39	150.02	372.10
% growth	96%	121%	168%	21%	32%

[Notes] Different thresholds for top earners: column (1) top 1% in overall distribution based on Census wage, (2) top 1% in overall distribution based on Piketty and Saez, 2003 (3) count of entertainer in top percentile, (4) 95th percentile of entertainer wage distribution, (5) 95th percentile of entertainer in 1940. Source: Data US Census and Piketty & Saez. Specification and sample same as baseline

Table 13: Effect of TV on Top Earner - Micro Data

	Probability in Top 1%			
	(1)	(2)	(3)	(4)
TV × Performance Entertainer	0.74	0.76	0.79	0.79
	(0.23)	(0.26)	(0.36)	(0.22)
TV × Interactive Leisure				-0.49
				(0.34)
TV × Drink & Dine				-0.65
				(0.48)
TV × Professional Services				0.32
				(0.21)
TV × Medics				-1.54
				(0.60)
TV × Engineer				-0.09
				(0.26)
TV × Manager				0.43
				(0.28)
Location & Occupation-Year FE	Yes	Yes	Yes	Yes
Demographics	–	Yes	–	Yes
Local labor market trends	–	–	Yes	–

[Notes] The outcome is a dummy that takes the value 100 if an individual is in the top 1% in the US distribution. Columns 1-3 are based on 83,748 individuals and column 4 on 3,438,002 individuals. Placebo occupations are non affected free time professions: drink & dining and active leisure and typical high pay professions: management, medicine, engineering, professional services (finance, accounting, law). The number of observations are 100308. Regressions use provided Census weights and cluster by local labor market.

Table 14: Effect of TV on Top Earner - State Level

	Share in Top 1%		
	(1)	(2)	(3)
Local TV station (1940)			-9.62
			(5.95)
Local TV station (1950)	20.94	20.18	-2.98
	(8.09)	(7.36)	(1.79)
Local TV station (1960)			-9.95
			(6.17)
Local TV station (1970)			-13.33
			(8.07)
Years	1940-1970	1916-1970	1916-1970
Observations	912	1008	1008

[Notes] Data US Census (1940-1970 and IRS in 1916. The regressor is the number of TV stations in 1950 in the state, allowing for time varying effects. In column 3 the omitted year is 1916. Standard errors are clustered at the state level.

Table 15: Earning Effect - triple diff

	Share in Top 1%		
	(1)	(2)	(3)
TV × Placebo Occupation	-0.41		
	(0.47)		
TV × Performance Entertainer	4.87	4.87	4.17
	(2.16)	(2.16)	(1.57)
TV × Interactive Leisure		-3.40	
		(1.29)	
TV × Drink & Dine		-3.80	
		(1.84)	
TV × Professional Services		5.23	
		(4.86)	
TV × Medics		-3.24	
		(1.52)	
TV × Engineer		-1.12	
		(1.23)	
TV × Manager		3.55	
		(2.21)	
Location & Occupation-Year FE	Yes	Yes	–
Pairwise Interaction: Location, Year, Occupation FE	–	–	Yes

[Notes] Data and specification are as in 1. Placebo occupations are non affected free time professions: drink & dining and active leisure and typical high pay professions: management, medicine, engineering, professional services (finance, accounting, law). The number of observations are 100,308.

Table 16: Quantile Effect of TV

	Wage Percentiles			
	99th	95th	75th	50th
Local TV station	260.3	85.00	22.33	19.13
	(92.23)	(3412.5)	(445.3)	(101.2)

[Notes] The reported coefficients are estimates using the quantile estimator for within group transformation developed in Powell (2016).

Table 17: Small Sample Performance of Pareto Shape Parameter Estimators

Estimator	sample 10%	local 5% sample
True	0.460	0.460
OLS	0.558	0.715
MLE	0.617	0.629
MLE (top code)	0.640	0.618
Close to cut-off	0.478	0.480

The true $1/\alpha$ is the value implied by the top 5% income share. The simulation draws samples from the entertainer wage distribution in the 1940 US full count Census. The samples are top coded at the 99th percentile of the distribution. Column 1 fits estimators on 10% samples dropping observations in the bottom half of the sample. Column 2 draws a smaller sample equivalent to a 5% sample of local labor markets. Estimates that imply an infinite mean are discarded ($\alpha < 1$)

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1662	Enrico Moretti Claudia Steinwender John Van Reenen	The Intellectual Spoils of War? Defense R&D, Productivity and International Spillovers
1661	Decio Coviello Andrea Ichino Nicola Persico	Measuring the Gains from Labor Specialization
1660	Nicolás González-Pampillón	Spillover Effects from New Housing Supply
1659	Walter Bossert Andrew E. Clark Conchita D'Ambrosio Anthony Lepinteur	Economic Insecurity and the Rise of the Right
1658	Paul Frijters Andrew E. Clark Christian Krekel Richard Layard	A Happy Choice: Wellbeing as the Goal of Government
1657	Philippe Aghion Antonin Bergeaud Matthieu Lequien Marc Melitz	The Heterogeneous Impact of Market Size on Innovation: Evidence from French Firm-Level Exports
1656	Clare Leaver Renata Lemos Daniela Scur	Measuring and Explaining Management in Schools: New Approaches Using Public Data
1655	Clément S. Bellet Jan-Emmanuel De Neve George Ward	Does Employee Happiness Have an Impact on Productivity?
1654	Matej Bajgar Giuseppe Berlingieri Sara Calligaris Chiara Criscuolo Jonathan Timmis	Industry Concentration in Europe and North America

1653	Andrés Barrios Fernandez	Should I Stay of Should I Go? Neighbors' Effects on University Enrollment
1652	Emanuel Ornelas Laura Puccio	Reopening Pandora's Box in Search of a WTO-Compatible Industrial Policy? The Brazil – Taxation Dispute
1651	Giulia Giupponi	When Income Effects are Large: Labor Supply Responses and the Value of Welfare Transfers
1650	Jonas Jessen Sophia Schmitz Sevrin Waights	Understanding Day Care Enrolment Gaps
1649	Stefano Federico Fadi Hassan Veronica Rappoport	Trade Shocks and Credit Reallocation
1648	Max Nathan	Does Light Touch Cluster Policy Work? Evaluating the Tech City Programme
1647	Stuart Campbell Lindsey Macmillan Richard Murphy Gill Wyness	Inequalities in Student to Course Match: Evidence from Linked Administrative Data
1646	Cong Peng	Does E-Commerce Reduce Traffic Congestion? Evidence from Alibaba Single Day Shopping Event
1645	Dan Andrews Chiara Criscuolo Peter N. Gal	The Best versus the Rest: Divergence across Firms during the Global Productivity Slowdown
1644	Christopher Cornwell Ian M. Schmutte Daniela Scur	Building a Productive Workforce: The Role of Structured Management Practices

The Centre for Economic Performance Publications Unit
Tel: +44 (0)20 7955 7673 Email info@cep.lse.ac.uk
Website: http://cep.lse.ac.uk Twitter: @CEP_LSE