Technical Change and Superstar Effects: Evidence from the Rollout of Television*

Felix Koenig[§]

December 2021

Abstract

Technical change that extends market scale can generate winner-take-all dynamics, with large income growth among top earners. I test this "superstar model" in the entertainer labor market, where the historic rollout of television creates a natural experiment in scale-related technological change. The resulting inequality changes are consistent with superstar theory: the launch of a local TV station skews the entertainer wage distribution sharply to the right, with the biggest impact at the very top of the distribution, while negatively impacting workers below the star level. The findings provide evidence of superstar effects and distinguish such effects from popular alternative models.

Keywords: *Superstar Effect, Inequality, Top Incomes, Technical Change*

JEL classification: J31, J23, O33, D31

*I thank the editor Amy Finkelstein and three excellent referees for their comments and ideas. I also 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, Brian Kovak, Camille Landais, Matt Lowe, Alan Manning, Ben Moll, Niclas Moneke, Markus Nagler, Alessandra Peters, Barbara Petrongolo, Steve Pischke, David Price, Arthur Seibold, Marco Tabellini, Lowell Taylor, Catherine Thomas, Joonas Tuhkuri, Anna Valero and John van Reenen as well as seminar participants at LSE, MIT, Goethe University, ZEW, RES Junior Symposium, IZA, EMCON, Ski and Labor, EEA Annual Congress and EDP Jamboree. I am grateful for financial support from the ESRC doctoral center during this project. [§]Carnegie Mellon University, Heinz College, Hamburg Hall, Pittsburgh, PA 15213, E-Mail: fkoenig@cmu.edu; Web: www.felixkoenig.com.

1 Introduction

In a celebrated 1981 article Sherwin Rosen argues that technical change can amplify inequality at the top of the wage distribution and generate extremely well-paid "superstar" earners. The driving force of such superstar effects is technological change that facilitate an increase in market scale. Rosen concludes that these technologies enable "many of the top practitioners to operate at a national or even international scale ... [and lead to] increasing concentration of income at the top" (Rosen 1981).[1] Superstar theory has been enormously influential, and has been used as the basis for modeling income inequality in a number of important settings.[2] Even so, there is little in the way of causal evidence for the theory, and Rosen's headline prediction about the impact of scale-related technologies on the earnings distribution remains to the best of my knowledge untested.[3]

This paper develops a test of the superstar model and uses a natural experiment to implement it. The paper studies the entertainment sector—arguably the most prominent field in which superstar effects are thought to arise—during the historic rollout of television.[4] Before the launch of television in the middle of the 20th century, successful entertainers typically had

[1]Early versions of the superstar theory appear in Tinbergen (1956), Sattinger (1975), and Sattinger (1979).

[2]See, for example, Terviö (2008), Gabaix and Landier (2008) and Gabaix, Landier, and Sauvagnat (2014) for CEOs; Garicano and Hubbard (2009) for lawyers; Kaplan and Rauh (2009) and Célérier and Vallée (2019) for finance professionals; Krueger (2005) and Krueger (2019) for entertainers; and Cook and Frank (1995), and Kaplan and Rauh (2013) for reviews in multiple sectors.

[3]A comparable state of affairs existed during the early development of theories and evidence concerning labor market effects of skill-biased technical change more broadly. Several observers, including Card and DiNardo (2002) and Lemieux (2006), stressed the need for clean identification to test theories of skill-biased technical change. Several subsequent studies indeed leveraged exogenous variation to implement such tests, e.g., Bartel, Ichniowski, and Shaw (2007); Akerman, Gaarder, and Mogstad (2015); Michaels and Graetz (2018); and Feigenbaum and Gross (2020).

[4]Many classic studies of superstar effects motivate their analysis with examples from the entertainment industry (see, e.g., Rosen 1981; Cook and Frank 1995; Krueger 2019).

live audiences of a few hundred individuals; after the launch of television audiences were an order of magnitude larger. In line with predicted superstar effects, I find that this shift resulted in disproportionate income gains for top entertainers, with much smaller gains for second tier stars and serious adverse effects for average talents.

The test uses a set of predictions of the superstar model that were at the heart of Rosen's original article. Existing theoretical work focuses on cross-sectional predictions of superstar-effects and compares the dispersion of talent to the dispersion of incomes. Unfortunately, though, it is challenging to test cross-sectional predictions because such tests require a credible cardinal measure of talent. Instead, I use the canonical model to derive predictions that focus on *changes* to inequality that can be tested without data on the talent distribution. These changes where the focus of Rosen's original argument and occur during periods of technical change, specifically when "scale related technical change" (SRTC) relaxes diseconomies of scale, thereby making large-scale production feasible.[5]

The model predicts that SRTC magnifies superstar effects and produces inequality changes that are different from other classic models of technical change. During SRTC, the most talented workers in the profession (the "superstars") attract an increased share of customers at the expense of lower-ranked talents. This process creates a few winners, with very high incomes, while reducing the total number of jobs. The right tail of the income distribution grows, and incomes become concentrated at the top, while employment and returns of lower-level talents decline.

The US government deployment plan of early television stations provides clean variation for a major SRTC, which facilitates a test of superstar model predictions. Entertainer audiences expanded via TV, and shows eventually reached a national audience. However, this transformation took place in

[5]Superstar effects also require imperfect substitutability of talent, as is typically the case in the entertainment sector.

stages.[6] Shows on early TV stations were broadcast via airwaves to the local population, and in this pioneering period technological constraints required TV shows to be filmed near the broadcast antennas. As a result, filming occurred simultaneously in multiple local labor markets, providing entertainers with a bigger platform in locations where stations were launched. Pioneering work on the US television rollout in Gentzkow (2006) and Gentzkow and Shapiro (2008) used the staggered rollout process and regulatory interruptions as a natural experiment to study impacts on television viewers. Building on this work, I study the effect of television on workers in the entertainment sector. The study uses a difference-in-differences (DiD) analysis across local labor markets during the staggered TV rollout, and leverages government rollout rules that led to quasi-random variation in the timing of TV launches.

The launch of a TV station skewed the entertainer income distribution to the right, with most of the skew happening in the very top tail. The fraction of entertainers with incomes that reach the top 1% of the US wage distribution doubled, with smaller increases at slightly lower income levels. Further down in the distribution such gains disappear, and lower-ranked talents lost out. The share of entertainers with mid-paid jobs declined, and the total employment of entertainers contracted by approximately 13%. In short, SRTC moved the entertainment industry toward a winner-take-all extreme, as predicted by superstar theory.

Two additional sources of variation help to strengthen the identification strategy. First, I use an unplanned interruption of the television deployment process. During the interruption, a group of places that were next in line for television had their permits blocked. Newly collected data on pending regulatory decisions allows me to track affected places.[7] Such places show

[6]Mass media (e.g., radio, newspapers and cinema) predates television and could reach a national audience. Television made additional entertainment formats scalable, and the local variation largely unfolds orthogonal to established media formats.

[7]Previous studies indirectly use this interruption period, but lacked the data to identify

no evidence of spurious shocks, and results from placebo tests support the assumption that the television deployment was exogenous to local demand conditions. Second, I exploit the staggered launch of TV stations combined with a subsequent decline of *local* filming. The advent of TV recording should remove local superstar effects of local stations, as recordings can be broadcast nationally. I confirm that the local treatment effect disappear after the emergence of TV recordings.

My study contributes to the literature on the impact of technical change on the labor market. Influential work has analyzed effects on the skill premium (Katz and Murphy 1992; Autor, Levy, and Murnane 2003; Acemoglu and Autor 2011; Autor 2014; Autor, Goldin, and Katz 2020) and on routine occupations (Autor, Katz, and Kearney 2008; Acemoglu and Autor 2011; Autor and Dorn 2013), and subsequent work has tested and confirmed these theories in natural experiments (see, e.g., Bartel, Ichniowski, and Shaw 2007; Akerman, Gaarder, and Mogstad 2015; Michaels and Graetz 2018; Feigenbaum and Gross 2020). My work is most similar to these latter studies; my distinctive contribution is to use a natural experiment to test whether SRTC has the effects predicted by superstar theory.

2 Superstar Effects and SRTC

The canonical superstar model, following Gabaix and Landier (2008) and Terviö (2008), features workers (actors) with heterogeneous talent (t) and employers (theaters) of varying size (s) and a production function where t and s are complements ($\frac{\partial Y}{\partial s \partial t} > 0$). For simplicity, assume t and s are Pareto distributed with shape parameters α and β respectively. The share of individuals with talent bigger than t, denoted by p_t, is therefore given by the CDF's complement $p_t = t^{-\frac{1}{\beta}}$ and similarly for s, $p_s = s^{-\frac{1}{\alpha}}$.

Turning to the revenue of a theater, I follow the literature on superstar

specific locations held up by the interruption.

effects and assume that each theater hires only one entertainer. The revenue of a theatre of size s that hires worker with talent t is $Y(s,t) = \pi(st)^{1/\phi}$, with output prices π and ϕ the degree of returns to scale in production which determines the cost of large-scale production.[8] SRTC changes parameter ϕ.

We can use this set-up to characterize the equilibrium wage distribution in this superstar economy. The share of workers with income above ω, denoted by p_ω, is (for derivations, see appendix A.1):

$$ln(p_\omega) = \gamma_0 - \gamma_1^\omega \phi, \tag{1}$$

with $\gamma_0 = \frac{\phi}{\alpha+\beta} ln(\frac{\beta\pi}{\alpha+\beta})$, and $\gamma_1^\omega \equiv \frac{ln(\omega)}{(\alpha+\beta)}$ functions of model parameters.

To formalize Rosen's notion of superstar effects, consider a SRTC that expands production scalability by decreasing ϕ. We can use equation (1) to evaluate the impact of such technical change on the wage distribution. Both terms of equation (1) are affected by ϕ. The first term affects all workers equally and captures the general change in marginal cost and marginal revenues. The gains, by contrast, are unequally distributed. These unequal effects are captured by γ_1^ω which varies with the wage level ω. In particular, γ_1^ω is bigger for larger ω and SRTC thus generates larger gains at the top of the distribution. The impact of such SRTC on inequality can be summarized in testable predictions about the income distribution (for proofs, see Appendix A.2):

Proposition. *In the superstar economy, SRTC leads to*

a) *Top wage growth: For sufficiently high income k, SRTC increases the share of workers above this extreme income level: $\triangle ln(p_\omega)|_{\omega>k} > 0$;*

b) *Fractal inequality: The effect of SRTC on the share of workers with income grater than ω (p_ω) is larger at higher income levels: $\triangle ln(p_\omega) > \triangle ln(p_{\omega'})$ if $\omega > \omega'$;*

[8] The results hold for broader functional form assumptions, as highlighted by Rosen 1981; Gabaix and Landier 2008; Terviö 2008.

c) Adverse effects for lesser talents: Employment at mid-pay levels declines; and

d) Employment loss: For a given outside option w^{res} and corresponding participation threshold $\bar{p}$, SRTC increases the participation threshold $\bar{p}$, i.e. $\frac{\partial \bar{p}}{\partial \phi} < 0$.

The impact of SRTC varies by income level and moves the labor market towards a winner-take-all setting. Propositions a) and b) highlight that SRTC disproportionally benefits the very top tail of the distribution—the superstars. Specifically, SRTC produces an increase in the share of workers with extremely high incomes (a) and widens the gap between the very top tail and slightly lower ranked talents (b). As stars serve a larger share of the overall market, incomes for lower ranked workers decline and the share of mid-paid jobs decreases (c). Finally, at the bottom end of the distribution, additional low-paid jobs emerge and in partial equilibrium with exit, workers will leave the industry (d).[9] These predictions are summarized in Figure 1a, while Figure 1b shows the corresponding empirical results that will be discussed below. At the top end of the distribution, the figure looks like an upward pointing hockey stick as the right tail of the distribution grows at the fastest rate (red markers indicate narrower wage ranges). By contrast, the share of workers with mid-paid jobs is reduced and more workers are in low paid positions.

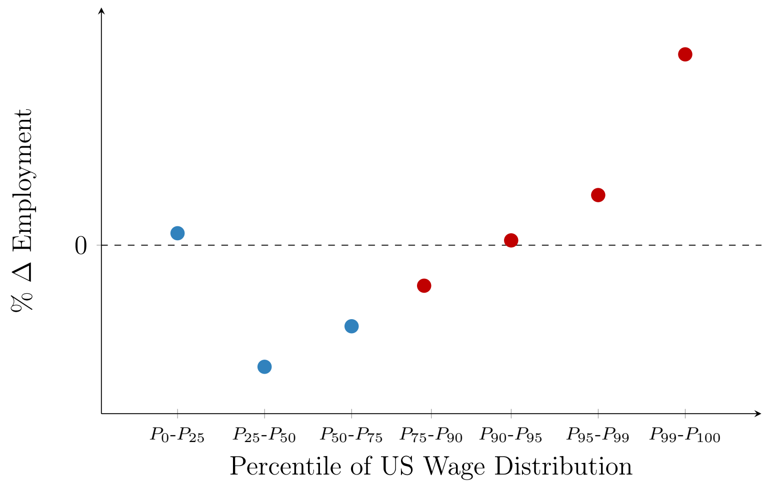
Superstar effects differ from alternative models of technical change in several ways. Canonical skill biased technical change, for example, features only two skill groups and thus produces little top income dispersion. Even extensions to SBTC models with greater type heterogeneity will struggle to generate top income inequality, particularly of the fractal nature described above. To replicate fractal inequality with SBTC, we would need to get rid

[9]The extend to which the low-paid sector grows, or exit occurs, depends on parameter assumptions.

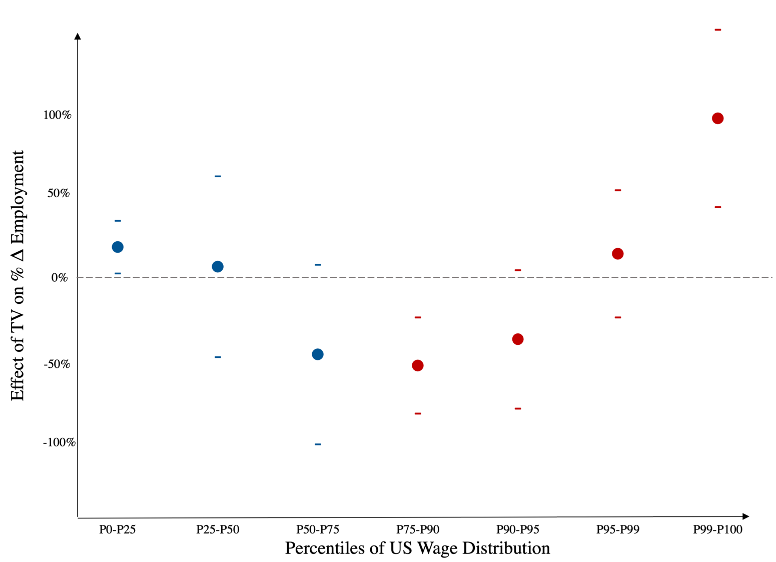
of the groups of perfectly substitutable workers and introduce imperfect substitution between workers. This is in principle feasible by taking the number of skill groups to infinity. Such an approach, however, is unattractive, as it introduces infinitely many parameters and makes the model impossible to falsify. The superstar economy instead provides a parsimonious and thus falsifiable model of income inequality. A second challenge for models with labor augmenting technical change is to generate real wage *and* employment losses (Caselli and Manning (2019)). The superstar framework produces such losses naturally, as shown by (c) and (d). Finally note that this benchmark superstar model is a closed economy model. Under plausible assumptions, the above inequality results do carry over to an open economy context where entertainment can be imported or exported from neighboring areas (see appendix A.3 for such an extension).

Figure 1: *SRTC Effect: Entertainer Employment at Different Wage Levels*

(a) Theoretical Predictions



(b) Empirical Difference in Difference Estimates



[Note] The figure shows the impact of SRTC on employment growth at different wage levels. Blue markers show effects at the bottom three quartiles and red markers at narrower ranges at the top. Panel A shows the theoretical predictions of SRTC, derived by differentiating equation 1 and parametrizing the result for a stylized shock that changes the scale parameter by factor 1.3 and the intercept by 0.2. The wage levels that correspond to P_{xx} are chosen from the US wage distribution to simplify comparison with the empirical results. Wages outside the range of previous support are grouped with the final bins to avoid undefined growth rates. Panel B shows empirical results from estimating equation 2, restricting $\beta_t = 0$ after the videotape launch in 1956. Each dot is a β coefficients from a separate with the outcome variable shown on the x-axis: P0-P25, for instance, refers to the share of entertainers with income between the 0th and 25th percentile of the US wage distribution. Dashes indicate 95% confidence intervals. Sources: US Census 1940–1970.

An alternative class of models has introduced task-specific technical change (for a summary, see Acemoglu and Autor (2011)). Such models have similarities to the superstar framework in that the latter also uses an assignment process to assign workers to tasks (or stages in our case). The task framework can produce real wage declines in response to labor augmenting technical progress by shifting workers into other tasks and increasing the supply of workers to such tasks. When it comes to top-income dispersion, task models face similar limitations to SBTC: Workers of equal skills are perfect substitutes and wage dispersion thus arises across skill groups only. We can generate a task model that is near isomorphic to a superstar model by letting the number of tasks and skill groups approach infinity.

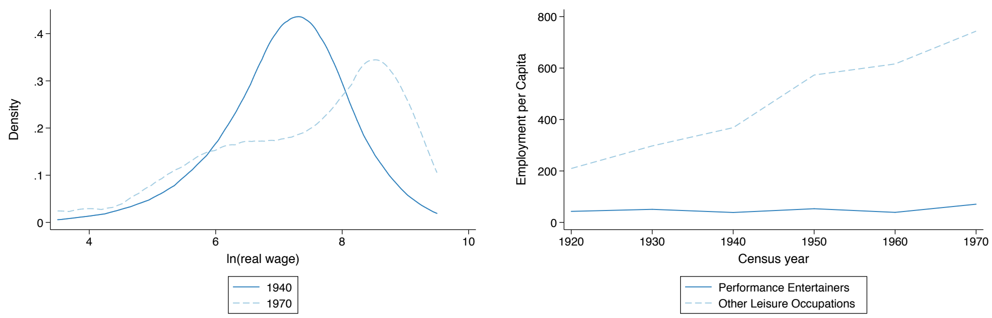
A remaining difference is how the two models conceptualize technical change. The task model studies the impact of factor augmenting shocks. To produce fractal inequality with such shocks, one has to assume that the technological shock is fractal itself, in the sense that technology boosts productivity most for the highest productivity workers. And while it is thus possible to generate fractal inequality, we would essentially assume the conclusion that we generate. In the superstar framework technical change (SRTC) affects a different parameter (the scale parameter) and the effect on labor demand at different skill levels arises endogenously, producing fractal inequality.

3 Empirical Test of Superstar Effects

I test the predictions of the superstar model by analyzing changes in the labor market for entertainers during the rollout of television. Television is a canonical case of SRTC, sharply expanding audience reach for entertainers. At the aggregate level, inequality in entertainment grew in line with superstar effects. Between 1940 and 1970, the entertainer wage distribution became more right skewed, and mid-paid positions were reduced (Fig 2a). At the same time, employment growth among entertainers lagged behind the rest of the leisure industry (Fig 2b).

Figure 2: *Change in Entertainment 1940–1970*

(a) *Entertainer Wage Distribution* (b) *Entertainer per Capita*



[Notes] Panel A shows the entertainment 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 (bars & restaurants and "other entertainment occupations"). The mean for performance entertainers is 49 and for other leisure occupations 468. Sources: US Population Census.

To estimate the causal impact of television, I exploit several features of the rollout process. First, technological and regulatory constraints largely confined pioneering TV stations to film live and locally to the broadcast antenna. And while some shows were relayed nationally, early transmission technologies led to poor image quality, which meant that such non-local shows were only used sparingly. The launch of an antenna thus produces SRTC for entertainers in the local area. The rollout of TV stations happened through a government deployment plan that was based on pre-determined local characteristics of the location (see Gentzkow (2006) and Gentzkow and Shapiro (2008)). I digitize the priority rankings, alongside the rules that went into designing them from "Television Digest" reports (for additional information on the data, see B.1).

A second source of variation is an unplanned interruption of the rollout in 1948. Several stations narrowly lost out on launches when the FCC discovered an error in their signal propagation model. This model was used to delineate interference-free signal catchment areas, but the error implied that signal interference occurred between neighboring stations. To avoid a worsening of the situation, the FCC put all licensing on hold and ordered a review of the model. I digitize information on these blocked locations, and use them for placebo tests in the analysis. The deployment resumed only 4 years later after extensive field work.

Finally, we also observe the decline of local TV filming. The invention of the videotape recorder in 1956 made local filming obsolete and shifted filming to the national level. We can use this period to test whether local stations stop generating local superstars. During this videotape era, we have to account for the emergence of national filming hubs, and regressions will include fixed effects for hubs in the post-videotape period. To avoid a potential endogenous control issue, I do not control for filming hubs directly but use a proxy for comparative advantages of a location as a filming hub. These proxies are based on a location's fixed characteristics, such as sunshine hours and landscapes, that largely drove location decisions. I quantify such predetermined factors using the share of movies filmed in the local labor market in 1920.

The baseline analysis uses a DiD design that compares local labor markets during the launch and subsequent decline of local filming. The analysis covers five occupations that commonly appeared on television (actors, athletes, dancers, musicians, and entertainers not elsewhere classified) and analyses the 722 US mainland local labor markets defined by Autor and Dorn (2013). Data on local entertainers come from the US decennial population Census and span the decades 1940-1970, while data on television exposure comes from hand-collected data on the location of television stations during the rollout (for additional information on the data, see B.1). The DiD estimation equation is:

$$Y_{mot} = \alpha_m + \delta_{ot} + \gamma X_{mt} + \beta_t TV_{mt} + \epsilon_{mot}. \tag{2}$$

where Y_{mot} measures labor market outcomes in labor market (m), year (t), occupation (o) (e.g., the share of entertainers in the top 1% of the wage US distribution), α_m and δ_{ot} are labor market and occupation-year fixed effects, and X_{mt} is a vector of control variables that includes the control for filming hubs of the post-videotape period. The treatment variable, TV_{mt}, is the number of local TV stations (for results with a TV dummy, see Appendix B.3.1). β_t captures the effect of local television stations, and this effect is allowed to change with the invention of the videotape in 1956. The standard errors ϵ_{mot} are clustered at the local labor market level so that running the analysis at the disaggregated level will not artificially lower standard errors. The full sample includes 722 local labor markets, 5 occupation groups and 4 Census years (2 for athletes, as the category is discontinued) and hence uses 13,718 observations and 722 clusters.

The first set of outcomes tracks changes in the entertainer earnings distribution. Following equation (1), I compute p_ω—the share of entertainers that reach income level ω—in each local labor market. I process the data in two ways. First, I deflate wage thresholds ω to make comparisons over time easier. The baseline deflator is the national wage growth at the corresponding wage percentile (results are robust to alternative deflators). With this adjustment, the outcome variable measures the rank position of entertainers in the US wage distribution of all workers. For example, a measure of extremely high incomes in entertainment is the share of local entertainers that reach the top 1% of the national US wage distributions.[10] A potential issue with these shares is that fluctuations in the denominator can generate spurious effects. To prevent this, I fix the denominator. The baseline results use the average employment count across labor markets as the denominator.[11] The

[10] Wages in the US Census are top-coded and the wage top code bites above the 99th percentile of the US distribution. We can thus identify all workers in the top 1% but cannot analyze more granular fractiles.

[11] To interpret the estimates as percentage point changes, I normalize by the average number of entertainers in *treated* labor markets.

results are, however, robust to alternative normalizations or using no such normalization (for robustness checks, see Appendix B.3.3).

The television variation has several advantages and addresses three "endogeneity challenges" which have made it difficult to obtain causal estimates of the effect of technological change on inequality. First, the government deployment process breaks the direct link between local economic conditions and television launches, and avoids the simultaneity problem that arises from ordinary, endogenous technology adoption.[12] Second, television is a SRTC that affects one market only—the provision of in-home broadcasting—and unlike other SRTCs (such as the internet) has little impact on other markets, which makes it a particularly clean case of SRTC. Finally, the rich variation from blocked stations and the timing of launches and removal enables us to distinguish the effect of television from simultaneous economy wide trends (such as deregulation, shifting norms, etc.).

4 Empirical Results

4.1 Changes in the Wage Distribution

As a starting point, I estimate the DiD equation (2) to test the superstar economy predictions outlined in Proposition 1. Proposition a) states that the share of entertainers at extreme income levels increases. An example of an entertainer who became famous through the launch of a local TV station is Korla Pandit, the star of "Korla Pandit's adventure in music" which aired multiple days a week on Los Angeles KTLA station. Korla was an eccentric organist who pretended to be Indian and performed exotic music in a light blue jewelry decorated turban. His show was a local favorite and he was voted "the local personality most deserving of national recognition" in a poll

[12] For a discussion of endogenous technical change see Acemoglu (1998), and for historical evidence see Beaudry, Doms, and Lewis (2010).

of the local LA *TV Guide* magazine.[13] To test for such effects formally, I study the impact of television on the share of entertainers among the top 1% highest-paid Americans by estimating equation (2) with a single time-invariant $\beta = \beta_{t<1956}$ that captures the average effect of stations before the launch of the videotape. The estimate and corresponding standard errors are shown on the far right of Figure 1b on page 9. A local TV station increases the share of extremely high paid entertainers by 4 percentage points on a baseline of 4 percentage points. A local TV station thus roughly doubles the share of local entertainers at these extreme income levels.

I next repeat the DiD estimation for slightly lower income, between the 95th and 99th percentiles. Figure 1b shows the coefficient next to the results of the previous regression. Strikingly, the effect of television on this income range is already substantially weaker. The point estimate shows a 50 percent increases in employment of entertainers in this range, an effect that is only half as big as the impact at the very top. Indeed, we cannot rule out that television had no effects on employment growth between the 95th and 99th percentiles. The polarization of the income distribution thus occurs predominantly in the very top tail. Television disproportionally benefit a small group of entertainers with the most extreme incomes, as superstar theory predicts.

Next, I test the effect of television on the rest of the distribution and continue to repeat the DiD analysis for lower income ranges. Figure 1b also plots the coefficients from these DiD regressions. The empirical estimates closely mirror the theoretical predictions in panel A. The right-most point reports the results on extreme income growth just discussed. Moving to the next-lower income range between the 90th and 95th percentiles, I find that television still has some positive effects but the effects are substantially smaller compared to the impact on at the very top. This confirms that the impact of SRTC diminish rapidly, even within the top tail of the distribution (confirming proposition b). Adverse consequences become visible at slightly

[13]The life story of Korla Pandit was recently turned into a film called "Korla."

lower income levels, and all income ranges from the 25th to the 90th percentile are negatively affected. Television thus hollows out the middle of the income distribution (proposition c). At the same time, the share of low paid entertainers increases. These results thus confirm the predictions of propositions a)-c).

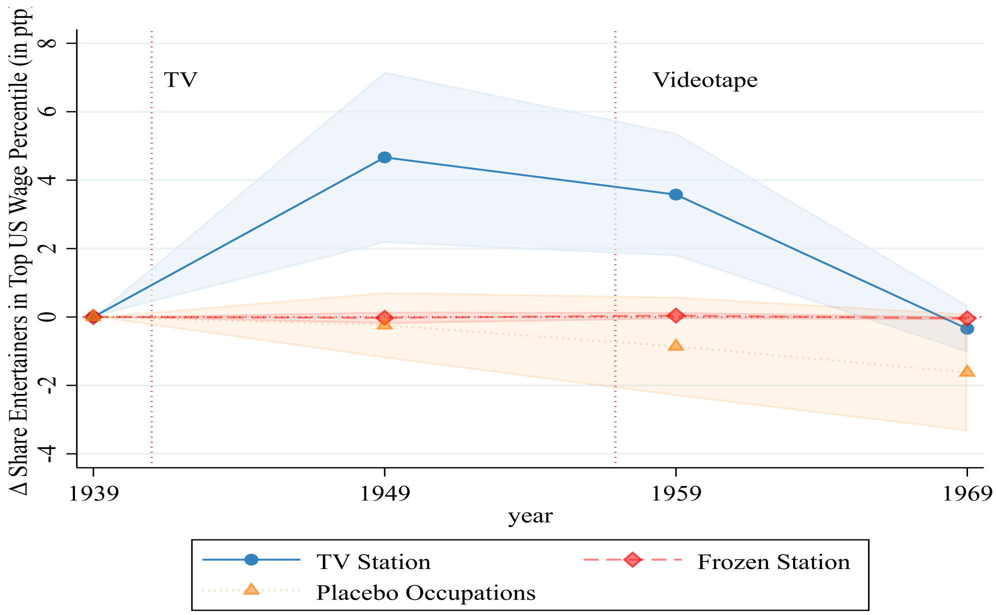
Alternative closely-related outcome variables are incomes at different percentiles of the wage distribution. Results for wage percentiles are reported in Appendix B.3.2. The predictions for wage percentiles can be obtained by re-arranging (1) and there is thus a one-to-one correspondence between wage results and the baseline results presented here. An advantage of the baseline results over quantile regressions is that the denominator of the outcome variable p_ω can be held fixed, as described above. This prevents spurious results from exits at the bottom end of the wage distribution which would artificially raise top wage percentiles.

4.2 Robustness Tests

The key identification assumption of the DiD is that TV launches are unrelated to spurious local shocks or trends that might similarly affect entertainer wages and employment. The television setting offers three compelling ways to probe this crucial assumption.

The first check relies on the *timing* of treatment effects. Specifically, we expect that effects which arise with the launch of a station should disappear after the demise of local filming. This is exactly what the blue line in Figure 3 shows. By 1970, after videotape's invention, the differences between treated and untreated locations revert back to their pre-treatment levels. Differences between treatment and control group only last for the duration of local television filming. There are thus no differential trends that create permanent differences between treatment and control areas. The career dent for local stars is again well illustrated by Korla Pandit. His attempt to acquire national fame failed and by the 1960s he reportedly returned to teaching piano lessons and playing various live venues, from mall openings, to concert halls.

Figure 3: *Dynamic Treatment Effect of TV on Changes in the Share of Entertainers in the Top US Wage Percentile*



[Note] Figure plots treatment coefficients from three dynamic DiD regressions. The specifications are dynamic versions of equation 2, with time-varying β_t. The outcome variable is the share of local entertainers (or placebo occupations) in the top percentile of the US wage distribution. Plot "TV Station" shows the coefficient on $TV_{m,t}$; plot "Placebo Occupations" is the same specification for the placebo occupations described in the text; plot "Frozen Station" uses blocked TV stations instead of $TV_{m,t}$ as treatment variable and uses untreated areas as control group. Vertical lines labelled *TV* and *Videotape* mark the beginning and end of local TV filming respectively. The area shaded in light blue marks the 95% confidence interval. Standard errors are clustered at the local labor market level.

Second, the interruption of the television rollout provides a powerful alternative test of spurious effects. Specifically, we can verify that places where station launches were unexpectedly blocked, i.e., places with "frozen stations," do not exhibit the same superstar effects that are observed in places that did launch. This is a placebo test for spurious local labor market shocks around the time of planned launches. The test is implemented in a dynamic DiD that compares the share of entertainers in the top 1% of US workers in untreated areas vs. areas that narrowly missed out on launches due to the rollout interruption. The red line in Figure 3 shows that blocked locations have no spurious changes before, after, or during the time of blocked launches. These results are precisely estimated and rule out even relatively small violations of the parallel trends assumption.

Third, since television only changed the production function of a handful of occupations, we can use other occupations for an additional placebo test. If TV assignment is orthogonal to local labor market conditions, we would expect that such placebo occupations would be unaffected. Different from the previous test, this test analyses spurious top income shocks in the same local labor market as the TV launch. An ideal placebo group would be able to pick up changes in top income in the local economy and I therefore use the main high-paying occupations as placebo groups (i.e., medics, engineers, managers and service professionals). As shown by the yellow line in Figure 3, there is indeed no effect on placebo occupations. Taken together, these parallel trend tests suggests that we are identifying the causal effect of television stations.[14]

The absence of spurious shocks is consistent with archival records of FCC decision making. Launch priority was based on pre-determined local characteristics (e.g., population size), and is thus unresponsive to local shocks. The empirical results confirm that these rollout rules were followed through in practice. Also notice that the results from this interruption experiment go beyond conventional pre-trend checks. Pre-trend checks focus on trends before the treatment, but with the blocked station experiment, we can additionally test for spurious shocks at the time of and after the planned TV launch date. For completeness, I also perform alternative robustness checks with conventional pre-trends (Online Appendix B.3.5) and triple differences (Online Appendix B.3.4) and both also find no spurious effects.

[14]We could perform triple-difference estimates by subtracting the results of either of these placebo tests from the baseline treatment effect estimates. Such a test would focus more directly on the difference in locations with approved and blocked stations. Such a test yields nearly identical results to the baseline DiD, as subtracting these placebo estimates from the baseline subtracts a value close to zero.

4.3 Employment Effects

The final prediction is employment loss implied by proposition d). These effects should occur in all areas that receive a TV signal. Since a television signal often extends beyond the local labor market where television filming takes place, more areas are affected by signal than by filming.[15] To account for this, I modify the estimation equation (2) and code all areas with signal as treated. The signal data comes from Fenton and Koenig (2020), and is a dummy with value one if signal is available in an area. The results show significant declines in employment (Table 1). The availability of signal reduces local entertainer employment by 13%, with similar results when controlling for demographics and local trends (column 2 and 3).

Panel B performs a placebo test with blocked stations and shows again that signal of blocked stations had no effect. A final robustness test focuses on differential pre-trends in treatment and control areas during the decade before the treatment by including a lead of the treatment in the regression. To perform this test, I expand the sample period backward by a decade.[16] The point estimate on the lead variable coefficient is small and insignificant and thus shows parallel pre-trends in the lead up to TV-signal (Panel C, column 4).

4.4 Alternative Effect Channels

For the interpretation of the results, it is useful to distinguish between two potential mechanisms: migration of entertainers and changing returns to talent. The Census asks individuals if they migrated recently and I use this information to test for entertainer migration responses, and find very small effects. The point estimates are negative and confidence intervals are tight

[15]This specification relaxes the closed economy assumption. With trade, the effects of a station launch spill into other labor markets. We explicitly model that employment losses may occur in labor markets that watch but do not produce television shows.

[16]This sample expansion is feasible since consistent employment data is available in the 1930 Census.

Table 1: *Effect of TV on Entertainer Employment*

	(1)	(2)	(3)	(4)
	Ln(Employment in Entertainment)			
		Panel A: Sample 1940—1970		
TV signal$_t$	-0.128	-0.114	-0.134	
	(0.061)	(0.061)	(0.063)	
		Panel B: Placebo Sample 1940—1970		
Placebo TV signal$_t$	0.053	0.044	0.053	
	(0.083)	(0.083)	(0.084)	
		Panel C: Sample 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)
No. of CZ cluster	722	722	722	722
Year–Occupation FE	Yes	Yes	Yes	Yes
CZ FE	Yes	Yes	Yes	Yes
Demographics	-	Yes	-	-
CZ level trends	-	-	Yes	-

[Note] The table shows the effect of television signal on local entertainer employment. Outcome variable *ln(Employment in Entertainment)* is the inverse hyperbolic sine of employment in entertainment. *TV signal* is a dummy that takes the value 1 if signal is available in a CZ, and *Placebo TV signal* if blocked stations would have brought TV signal. Column 2 controls for median age, % female, % minority, population density, and trends for urban areas. Column 3 controls for a separate linear trend for each CZ. Subscript *t+1* refers to the lead of the treatment variable. Panel A & B include 13,718 CZ-year-occupation observations and Panel C 17,328 observations. Standard errors are reported in brackets and are clustered at the local labor market level. Sources: TV signal from Fenton and Koenig (2020) and labor market data from US Census 1930–1970.

(Table 2, Panel A). Migration thus appears to contribute little to the results. A potential explanation for the limited mobility response is that early shows tended to focus on local events, following the tradition of vaudeville, and thus did not translate easily to other locations. We can use the mobility estimates to bound the impact of migration. The central estimates suggest that mobility plays next to no role in the results; even at the upper bound of plausible values, the migration channel can only explain a quarter of the total effect.

A related concern is commuting across local labor market boundaries. Such behavior would downward bias the estimates by spreading the impact of local shocks beyond the boundary defined by commuting zones. Commuting is arguably easiest between neighboring areas, and we can thus alleviate the impact on the results by excluding areas that are adjacent to television launch locations from the analysis. Results that exclude such neighboring areas show very similar effects to the baseline, indicating that commuting plays a minor role in these findings (Table 2, Panel B).

A further potential concern is that television changed the type of talent required of entertainers. As a result, television may have changed the distribution of talent, rather than returns to talent. The talent distribution is not directly observable but historic description of television recruiting at the time suggests that the types of talent remained the same. Early television relied heavily on established show formats and often broadcast vaudeville shows (for an overview see, Murray 1999). *Variety* magazine reported on the tensions that resulted from television poaching stars from established shows: "Criticism is being advanced in the trade that television so far has not kept its promise of developing its own talent." The television industry responded to this criticism and actively encouraged poaching, arguing that "stars are not going to be made by television. Television is going to be made by stars. So—let's go out and get them!"[17]

[17] See, respectively, Bob Stahl, "Where's that New TV Talent? Medium Scorned for it's

Table 2: *TV and Migration Between Labor Markets*

	(1)	(2)	(3)
		Panel A: Share Entertainers who Migrated	
Local TV stations	-0.014	-0.017	-0.010
	(0.015)	(0.015)	(0.020)
Sample	full	full	full
No. of CZ cluster	722	722	722
		Panel B: Entertainer among Top 1% of US Earners (excluding CZs adjacent to treated CZs)	
Local TV stations	4.30	4.46	6.16
	(1.31)	(1.30)	(2.27)
Sample	no adjacent CZ	no adjacent CZ	no adjacent CZ
No. of CZ cluster	568	568	568
Year–Occupation FE	Yes	Yes	Yes
CZ FE	Yes	Yes	Yes
Demographics	–	Yes	–
CZ level trends	–	–	Yes

[Note] The Table tests the effect of local TV launches on entertainer migration. Outcomes: Panel A, the fraction of entertainers who moved; Panel B, share of entertainers among the top 1% of the US wage distribution, excluding CZs that share a border with treated CZs. All regressions control for commuting zone (CZ), occupation specific time fixed effects and local filming cost in years after the invention of the videotape. Entertainers are actors, athletes, dancers, entertainers not elsewhere classified, musicians. Column 2 controls for median age & income, % female, % minority, population density, and trends for urban areas. Column 3 controls for a separate linear trend for each CZ. Sample: 13,718 observations and 722 CZs. Demographic data is missing for one CZ in 1940 and thus reduces the sample in column 2. The exclusion of CZs in Panel B reduce the sample to 10,792 observations and 568 CZs. Observations are weighted by local labor market population. Standard errors are reported in brackets and are clustered at the local labor market level. Source: US Census 1940-1970.

As a robustness test, we can probe the assumption of a stable talent distribution indirectly. This assumption implies that people maintain their rank in the distribution, while changes in the distribution would imply leapfrogging of individuals. To investigate leapfrogging, I digitize "who is who" lists of 1950 TV stars and merge individuals in these lists to their 1940 Census records to build a small panel. The 1940 earnings data enables me to analyze earning ranks of later TV stars in the pre-TV period. The panel data shows that in 1940, before television, roughy 60% of the later TV stars where already in the top wage decile and less than 5% were in the bottom decile. Eventual television stars were thus already disproportionally high-paid before television and television did not generate major leap-frogging (for data details, see Appendix B.3.6). Both historic sources and the data thus suggest that early television stations targeted the same talents who were successful in traditional entertainment.

5 Conclusion

It has been forty years since Sherwin Rosen (1981) presented his elegant superstar theory. In this influential work, Rosen shows how scale related technological change can serve as a driving force in the generation of income inequality, particularly at the top end of the income distribution.

This paper provides the first direct test of this theory, using a simple natural experiment, and finds clear evidence that scale related technological change can generate superstar effects, including income concentration at the top. The basis for the test is the increase in the market reach of entertainers that arose during the staggered introduction of television. The launch of a TV station increased audiences of star entertainers and created high paid superstar entertainers. The income distribution skewed to the right, with

Laxity," *Variety*. 26 Oct. 1949:1, and "Video Needs Comedy: Tele-viewers Prefer Variety Show," *Television World*. 24 May 1948:3.

escalating effects as we move up towards the top of the wage distribution. At the same time, the middle of the distribution is hollowed out as the share of entertainers with average incomes declines significantly, and many lesser stars lost their jobs.

Technical change is, however, unlikely to generate superstar effects in all sectors of the economy. Superstar effects arise only in sectors where talent is heterogeneous and unique; we expect superstars to be less important (or not important at all) in settings where individual-level talent is highly substitutable. The production technology thus plays an important role, and not all scale related technologies lead to superstar effects. An interesting avenue for future research is to explore which other sectors meet the conditions for superstar effects, and to quantify how the magnitude of these effects vary across different sectors of the economy.

A broad literature has recognized that a better understanding of superstar effects is important not only from a scientific standpoint, but also for policy decisions. Top earners are one of the main sources of tax revenue, and recent research shows that superstar effects could have substantial effects on the optimal level and progressivity of taxes (Scheuer and Werning 2017). Moreover, superstar effects might influence the potential benefits from policies that reduce economic concentration, and may explain economic divergence between regions (Eckert, Ganapati, and Walsh 2019). Finally, it is useful to remember that the full social welfare implications of SRTC include not only effects on the distribution of income—potentially via superstar effects—but also on the value attached to greater equality in access to goods and services (Acemoglu, Laibson, and List 2014). Scale related technological change is shaping many sectors of the economy, and it is thus important that we improve our understanding of the far-reaching economic consequences.

References

Acemoglu, Daron (1998). "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality." *The Quarterly Journal of Economics* 113.4, pp. 1055–1089.

Acemoglu, Daron and David H. Autor (2011). "Skills, Tasks and Technologies: Implications for Employment and Earnings." *Handbook of Labor Economics* 4.PART B, pp. 1043–1171.

Acemoglu, Daron, David Laibson, and John A. List (2014). "Equalizing superstars: The internet and the democratization of education." *American Economic Review* 104.5, pp. 523–527.

Akerman, Anders, Ingvil Gaarder, and Magne Mogstad (2015). "The Skill Complementarity of Broadband Internet." *Quarterly Journal of Economics* 130.4, pp. 1781–1824.

Autor, David H. (2014). "Skills, education, and the rise of earnings inequality among the "other 99 percent"." *Science* 344.6186, pp. 843–851.

Autor, David H. and David Dorn (2013). "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market." *American Economic Review* 103.5, pp. 1553–97.

Autor, David H., Claudia Goldin, and Lawrence F. Katz (2020). "Extending the Race between Education and Technology." *American Economic Review (papers and proceedings)*, pp. 49–54.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney (2008). "Trends in U.S. Wage Inequality: Revising the Revisionists." *Review of Economics and Statistics* 90.May, pp. 300–323.

Autor, David H., F. Levy, and Richard Murnane (2003). "The Skill Content of Recent Technological Change: An empirical exploration." *Quarterly journal of economics* 118.4, pp. 1279–1333.

Bartel, Ann, Casey Ichniowski, and Kathryn Shaw (2007). "How Does Information Technology Really Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement and Worker Skills." *The Quarterly Journal of Economics* 122.4, pp. 1721–1752.

Beaudry, Paul, Mark Doms, and Ethan Lewis (2010). "Should the personal computer be considered a technological revolution? Evidence from U.S. metropolitan areas." *Journal of Political Economy* 118.5, pp. 988–1036.

Card, David and John DiNardo (2002). "Skill Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles." *Journal of Labor Economics* 20.4, pp. 733–783.

Caselli, Francesco and Alan Manning (2019). "Robot Arithmetic: New Technology And Wages." *American Economic Review: Insights* 1.1, pp. 1–12.

Célérier, Claire and Boris Vallée (2019). "Returns to Talent and the Finance Wage Premium." *The Review of Financial Studies* 32.10, pp. 4005–4040.

Cook, Philip J. and Robert Frank (1995). *The Winner-Take-All Society*, p. 288.

Eckert, Fabian, Sharat Ganapati, and Conor Walsh (2019). "Skilled Tradable Services: The Transformation of U.S. High-Skill Labor Markets." *Working Paper*.

Feigenbaum, James and Daniel P Gross (2020). "Automation and the Fate of Young Workers." *NBER Working Paper Series*.

Fenton, George and Felix Koenig (2020). "Labor Supply and Innovation in Entertainment: Evidence from TV." *Working paper*.

Gabaix, Xavier and Augustin Landier (2008). "Why Has CEO Pay Increased So Much?" *Quarterly Journal of Economics* 123.1, pp. 49–100.

Gabaix, Xavier, Augustin Landier, and Julien Sauvagnat (2014). "CEO pay and firm size: An update after the crisis." *Economic Journal* 124.574, pp. 40–59.

Garicano, Luis and Thomas N. Hubbard (2009). "Specialization, firms, and markets: The division of labor within and between law firms." *Journal of Law, Economics, and Organization* 25.2, pp. 339–371.

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

Gentzkow, Matthew and Jesse Shapiro (2008). “Preschool Television Viewing and Adolescent Test Scores: Historical Evidence from the Coleman Study.” *Quarterly Journal of Economics* 123.February, pp. 279–323.

Kaplan, Steven N. and Joshua Rauh (2009). “Wall Street and Main Street: What Contributes to the Rise in the Highest Incomes?” *Review of Financial Studies* 23.3, pp. 1004–1050.

Kaplan, Steven N. and Joshua Rauh (2013). “It’s the Market: The Broad-Based Rise in the Return to Top Talent.” *Journal of Economic Perspectives* 27.3, pp. 35–56.

Katz, Lawrence F. and Kevin M. Murphy (1992). “Changes in Relative Wages, 1963-1987: Supply and Demand Factors.” *The Quarterly Journal of Economics* 107.1, pp. 35–78.

Krueger, Alan B. (2005). “The Economics of Real Superstars: The Market for Rock Concerts in the Material World.” *Journal of Labor Economics* 23.1, pp. 1–30.

Krueger, Alan B. (2019). *Rockonomics: What the Music Industry Can Teach Us About Economics and Life*. Crown, p. 336.

Lemieux, Thomas (2006). “Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?” *American Economic Review* 96.3, pp. 461–498.

Michaels, Guy and Georg Graetz (2018). “Robots at Work.” *Review of Economics and Statistics* 100.5, pp. 753–768.

Murray, Susan (1999). “"Hitch Your Antenna to the Stars!" Early Television and the Renegotiation of Broadcast Stardom.” PhD thesis. University of Texas at Austin.

Rosen, Sherwin (1981). “The Economics of Superstars.” *American Economic Review* 71.5, pp. 845–858.

Sattinger, Michael (1975). “Comparative advantage and the distributions of earnings and abilities.” *Econometrica* 43.3, pp. 455–468.

Sattinger, Michael (1979). “Differential Rents and the Distribution of Earnings.” *Oxford Economic Papers* 31.1, pp. 60–71.

Scheuer, Florian and Ivan Werning (2017). “The Taxation of Superstars.” *Quarterly Journal of Economics* 132.1, pp. 211–270.

Terviö, Marko (2008). “The difference that CEOs make: An assignment model approach.” *American Economic Review* 98.3, pp. 642–668.

Tinbergen, Jan (1956). “On the Theory of Income Distribution.” *Weltwirtschaftliches Archiv* 77.2, pp. 155–175.

ONLINE APPENDIX

A APPENDIX: Derivations

A.1 Equilibrium of the Superstar Economy

Each firm maximizes profits by hiring a worker with talent t, taking its own firm characteristic as given. The firm problem is therefore given by

$$max_t Y(s_i, t) - w(t),$$

where $w(t)$ is the wage for a worker with talent t. The equilibrium is characterized by the incentive compatibility condition, the participation condition, the assignment function of workers to firms, and market clearing.

The optimal assignment $\sigma(S_i) = t$ matches the best actor with the biggest theater. This PAM results follows from the comparative advantage assumption $\frac{\partial Y}{\partial t \partial S} > 0$, which implies better actors have a comparative advantage in bigger theaters. PAM guarantees that the percentiles of talent and size distribution are the same for a matched pair $p_s = p_t$. Moreover, since wages correspond to worker productivity, the percentile in the talent distribution corresponds to the percentile in the wage distribution $p_t = p_w$. Since the equilibrium is competitive, the optimal assignment is also the market outcome and hence the first equilibrium condition.

Incentive compatibility guarantees that for each firm i the optimal worker p meets,

$$Y(s_i, t) - w(t) \geq Y(s_i, t') - w(t') \quad \forall \ t' \epsilon [\underline{t}, \overline{t}]. \tag{3}$$

The number of incentive compatibility (IC) constraints can be reduced substantially. If the IC holds for the adjacent t' all the other ICs will hold as well. We can therefore focus on the percentiles just above and below t. The IC for the adjacent $t' = t + \epsilon$ can be further simplified if Y is differentiable in t. Divide equation 3 by ϵ and let $\epsilon \rightarrow 0$.

$$\frac{w(t) - w(t+\epsilon)}{\epsilon} \leq \frac{Y(s_i, t) - Y(s_i, t+\epsilon)}{\epsilon}$$

$$\frac{\partial w}{\partial t} = \frac{\partial Y(S_i, t)}{\partial t}. \tag{4}$$

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

I extend the model and allow for entry and exit. This gives rise to a fourth equilibrium object, the participation threshold $\bar{p}$, which is defined by the participation constraints (PC). Denote the reservation wage of workers $w^{res}$and the reservation profits ψ^{res} and hence the PC condition is

$$Y(s_i, t) - w(p) \geq \psi^{res} \quad \forall \;\; p\epsilon[\bar{p}, 1] \tag{5}$$

$$w(p) \geq w^{res} \quad \forall \;\; p\epsilon[\bar{p}, 1]. \tag{6}$$

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}.

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

Using the functional form assumptions in the text, we can rewrite (4) as

$$\frac{\partial w}{\partial t} = \frac{\pi}{\phi} s^{\frac{1}{\phi}} t^{\frac{1}{\phi}-1} = \frac{\pi}{\phi} t^{\frac{1}{\xi}-1}, \tag{7}$$

where $\xi = \frac{\phi}{\alpha+\beta}$, the last equality uses the size distribution and $p_s = p_t = p^w$.

Integrating and normalizing $w(\underline{t}) = 0$ gives the wage:

$$w(t) = \int_{\underline{t}}^{t} \frac{\partial w}{\partial t} = \frac{\pi\beta}{\alpha+\beta} t^{-1/(\xi\beta)} = \frac{\pi\beta}{\alpha+\beta}[p_w]^{-1/\xi}. \tag{8}$$

taking logs, evaluating w at ω and re-arranging yields equation 1:

$$ln(p_\omega) = \gamma_0 - \gamma_1^\omega \phi.$$

A.2 Proof of Proposition 1

This section derives the four parts of the Proposition in the text.

Part a. Compute the employment share that pays above ω (denoted by $ln(p^\omega)$) before and after SRTC by evaluating equation 1 at the two values of $\phi, \tilde{\phi}$ respectively before and after SRTC:

$$\Delta ln(p^\omega) = \tilde{\gamma_0} - \gamma_0 + \gamma_1^\omega(\phi - \tilde{\phi}).$$

This captures the change in $(ln(p^\omega))$. When $\omega \to \infty$, then $\gamma_1^\omega \to \infty$ and since SRTC implies $\phi > \tilde{\phi}$, this implies that the right hand side is positive. SRTC therefore increases the share of workers with extremely high incomes.

Part b. $\Delta ln(p^\omega)$ is bigger at higher income levels since γ_1^ω increases in ω: $\partial\gamma^\omega/\partial\omega = \frac{\omega}{\alpha+\beta} > 0$. The impact of SRTC is thus greater at higher income levels. Moreover, even the second derivative is positive, implying that the rate of increase also grows at higher income levels. In short, the right tail of the distribution gains disproportionally.

Part c. Define a mid-income workers as having a wage between w & w' and denote the share of mid-paid entertainers by M. This share can be derived using equation 1:

$$M = p(w) - p(w') = (\frac{\beta\pi}{\alpha+\beta})^\xi[w^{-\xi} - w'^{-\xi}].$$

Differentiating with respect to ϕ gives the impact of SRTC: $\partial M/\partial\phi = -\varepsilon_D\kappa +$

$(\partial M/\partial\xi)/(\alpha+\beta)$, where ε_D is the elasticity of inverse demand and $\kappa = \frac{\xi}{\phi}(\frac{\beta\pi}{\alpha+\beta})^{\xi}[w^{-\xi} - w'^{-\xi}]$. Mid-income jobs will decline when $\partial M/\partial\phi < 0$, which occurs when demand is sufficiently inelastic (i.e., if the elasticity of the inverse demand curve is $\varepsilon_D > \frac{\partial M/\partial\xi}{(\alpha+\beta)\kappa}$).[18] Note, however, that the previous equation only holds for wages that are in the support of the income distribution both before and after SRTC. Given that the wage distribution spreads out with SRTC, we may reach wage levels that were previously unattained and thus violate this condition. In such wage ranges, the growth rate is undefined. The share of entertainers in the baseline period is 0 and to compute a growth rate we would have to divide by 0. To get around this, I group newly emerging pay ranges together with the nearest wage that occurred before SRTC. In that case, employment shares at the extremes of the distribution increase unambiguously, and as a result we may see growth in low-paid employment.

Part d. In the model with entry and exit the participation constraint (PC) ensures that the marginal participant ($\bar{p}$) is indifferent between working and the outside option (w^{res}) and the marginal employer breaks even:

$$w(\bar{p}) = w^{res},$$

$$Y(\sigma(\bar{p}), \bar{p}) = w(\bar{p}).$$

A period of SRTC is such case that decreases $Y(\sigma(\bar{p}), \bar{p})$ by reducing π. To reach equilibrium, $\bar{p}$ has to adjust. Recall that low p implies a high level of talent and hence $dY(\sigma(\bar{p}), \bar{p})/d\bar{p} < 0$. The SRTC induced fall in Y therefor results in a in lower $\bar{p}$, which confirms Proposition (d).

[18]Notice that if M declines for an income range w to w', it will also decline for all lower income ranges. This follows since $\frac{\partial M/\partial\xi}{(\alpha+\beta)\kappa}$ is larger at higher values of w and therefore the elasticity condition will hold for lower wage ranges if it holds at M. The result that $\frac{\partial M/\partial\xi}{(\alpha+\beta)\kappa}$ increases with income follows because κ increases with income at a rate proportional to $[w^{-\xi} - w'^{-\xi}]$, while $\partial M/\partial\xi$ increases at a faster rate, proportional to $[w^{-\xi} - w'^{-\xi}] + [w^{-\xi}(ln(w) - 1) - w'^{-\xi}(ln(w') - 1)] > [w^{-\xi} - w'^{-\xi}]$.

A.3 Open Economy Extension

The above model is a closed economy model. In an open economy, entertainers can exported their services outside their local labor market. For example, local television stars may draw additional audiences from neighboring areas within the signal reach of the television antenna. Conversely, demand for local entertainers may decline in locations where television was not filmed but where shows from a television station from a neighboring area could be watched. To assess the impact of such a case, consider the effect of a station launch in m on entertainer revenues in a neighboring local labor market k. Assuming that local entertainer talent and theater capacity is fixed, the impact of an SRTC is:

$$\frac{\partial Y_k}{\partial \phi_m} = (s_k t_k)^{\phi_k} \partial \pi_k / \partial \phi_m.$$

The decline in demand affects local entertainers through a drop in the price for a quality adjusted unit of entertainment (π). As a result, the revenue from shows declines and some entertainers in k leave entertainment. Intuitively, a station launch in m produces employment losses in all labor markets reached by its signal.

Some individuals from k now consume entertainment originated in m and the corresponding export effect increases demand for local entertainers in m. This increase in demand increases the market clearing price for entertainment (π_m). This export effect on π_m goes in the opposite direction to the direct effect of television. The direct effect reduces the local marginal cost of providing entertainment and thus π_m. As long as the trade effect does not overturn the baseline effect of television, all the above predictions of the superstar economy still hold in the open economy.

One might contest that the impact of imports is better modeled as an increase in demand for *specific* entertainers, rather than as an general increase in demand for entertainment. Whether this a distinction matters depends on the market structure. In a market with a single market clearing price for entertainment, the incidence of a demand shocks does not matter. In this case, a demand shock for one entertainer affects the equilibrium price for entertainment and spreads out across all entertainers. In the superstar setting, there is no such single market price for talent units and different firms, instead, pay different rates for talent units. To analyze how the dispersion of talent prices is affected by trade, recall that the price for a talent unit is $w'(t_i) = \frac{\pi}{\phi} t_i^{\frac{1}{\xi}-1}$. As before, we assume that individual characteristics t_i are unchanged by television. The impact of SRTC on the price of the marginal talent unit of individual i in m is:

$$\frac{\partial w'(t_i)}{\partial \phi_m} = -\frac{w'(t_i)}{\phi_m} + \frac{\partial \pi}{\partial \phi_m}\frac{w'(t_i)}{\pi} + \frac{(\alpha+\beta)\cdot\beta}{\phi_m} ln(p_i)\frac{w'(t_i)}{\phi_m} = \frac{w'(t_i)}{\phi_m}[\epsilon_\pi - 1 + \frac{(\alpha+\beta)\cdot\beta}{\phi_m} ln(p_i)]$$

where ϵ_π is the elasticity of entertainment demand in market m. Re-arranging we can show that the elasticity of the talent price to SRTC ($\epsilon_{w'_i}$) is:

$$\epsilon_{w'_i} = \frac{(\alpha+\beta)\cdot\beta}{\phi_m} ln(p_i) - [1 - \epsilon_\pi]$$

This is the same expression as in the closed economy. The difference to the closed economy is that trade makes demand is more elastic, which means ϵ_π is larger. Trade thus changes the last term in the above expression. Note that the inequality results of SRTC are captured by the first term, which is unaffected by trade. The second term does not vary with i and therefore affects elasticities at all talent levels in the same way. In other words, trade has a level effect but otherwise stretches or compresses the wage distribution proportionally. As a result, the direction of the inequality predictions in proposition 1 are unchanged.[19] The intuition for this result is that trade affects the level of demand, but the *distribution* of demand across workers is

[19] This intuition breaks if television alters the distribution of s or t. If the distributions of s or t became more fat tailed, this further amplifies the inequality results.

pinned down by the production function and the scarcity of inputs.

An exception is the prediction of employment losses (prediction d), which is sensitive to assumptions about trade. The level of employment depends on the level of demand. Market m may experience no employment loss if exports to market k raise local entertainment demand in m enough and offset the decline in demand for lower ranking entertainers. In this case, the negative employment effects would be "off-shored" to market k. The derivation in section A.2 rules such cases out by imposing a restriction on the demand elasticity.

B APPENDIX: Empirics

B.1 Data Sources and Construction

Television Data

Data on the TV rollout is documented in publications of the FCC. The FCC decided how to prioritize areas during the TV rollout. I digitize the location of the approved launches. The data on TV launches is published in the annual *Television Yearbooks* and I collect this information and identify the CZ of each TV launch.[20] For TV signal, I use data from (Fenton and Koenig 2020) which compute signal catchment areas of historic TV stations. To compute similar signal reach for stations that were blocked, I additional collect records on the technical features of planned antennas. These details were recorded by the FCC to compute transmission areas and potential signal interference. I use this data to reconstruct the signal of TV stations that narrowly missed out on launches. The relevant FCC records are published as part of the *TV Digest* 1949.

To match TV signal exposure to the Census, I map county-level TV signal information onto geographic units available in the Census. The geographic

[20] Called *TV Digest* in earlier years.

match uses the boundary shapefiles provided by the National Historical Geographic Information System (NHGIS) (Manson et al. 2017).

Outcome variables The main outcome variable is the rank of local entertainers in the US income distribution. Consider the share of local entertainers that reaches the top 1% of the US income distribution. This takes value 0 when no entertainer earns such extreme wages and value 100 in a winner-takes-all market with a single superstar entertainer.[21] The share in market m at time t is:

$$p_{m,t}^{\omega^{99}} = \frac{\sum_{i\epsilon I} E_{i,m,t}}{\overline{E_t}}, \tag{9}$$

where E is a dummy that takes the value 1 for entertainer occupations and I is the set of workers in the top 1% of the US wage distribution. The wage top code bites above the 99th percentile of the US distribution and we can thus identify all workers in the top 1%.[22] A potential issue with these shares is that fluctuations in the denominator can generate spurious effects. To prevent this, I use the number of entertainers in the average labor market ($\bar{E}_t$) as denominator instead of local labor market counts.[23] As an alternative approach, I compute per capita counts which use the local population as the denominator. These measure map directly into the predictions presented in the text and measures how the top tail of the entertainer distribution stretches out relative to the US distribution. We can naturally extend the analysis to other percentiles and study where entertainers rank in the US distribution for all income ranges. Finally, I also compute the wage at the

[21]This metric is similar in spirit to Chetty et al (2017) who also study ranks of local workers in the national distribution. The authors highlight that such ranks have advantages over income levels for comparisons over longer time periods.

[22]The relevant top 1% thresholds are: 7,555 8,050 11,859 16,247 in 1950 USD for 1940, 1950, 1960 and 1970 respectively.

[23]To interpret the estimates as percentage point changes, I normalize by the average number of entertainers in *treated* labor markets. Note that this normalization also implies that $p_{m,t}^{\omega^{99}}$ can in principle be bigger than 100. This approach codes areas without local entertainers – for instance areas where television displaced all local entertainers – as 0. Robustness checks without the normalization show similar results (see Appendix B.3.3).

top percentile of the local entertainer distribution and top income shares of local entertainers.

Census Data Processing

Local labor markets The analysis defines a local labor market as a commuting zone (CZ). A labor market comprises an urban center and the surrounding belt of commuters. The CZs fully cover the mainland US. The regions are delineated by minimizing flows across boundaries and maximizing flows within labor markets, and 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 CZs (Autor and Dorn 2013). I use these crosswalks for the 1950 and 1970 data and build additional crosswalks for the remaining years. For each Census, I use historical maps for the smallest available location breakdown. I map the publicly available Census location identifiers into a CZ. 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, using mini public use microdata areas (mini-PUMAs). To crosswalk the 1940 data, I use maps that define boundaries of the identified areas. In geographic information system (GIS) software I compute the overlap of 1940 counties and 1990 CZs. In most cases counties fall into a single CZ. A handful of counties are split between CZs. For cases where more than 3% of the area falls into another CZ, I construct a weight that assigns an observation to both CZs. The two observations are given weights so that together they 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 emerges only in 1969 as a merger of Ormsby County and Carson City, 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.

Worker data Data is provided by the Integrated Public Use Microdata Files (IPUMS, Ruggles et al. 2017) of the US decennial Census from 1930 to 1970 (excluding Hawaii and Alaska). 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. This limits the potential to expand the sample. During the sample period most variables remain unchanged, and where changes occurred, IPUMS has aimed to provide consistent measures. For each of the years, I use the largest publicly available sample with granular spatial data; before 1950, data on the full population is available, and I use samples for recent years. In 1970 the biggest available dataset combines data from Form 1 and Form 2 metro samples. The data cover 722 CZs that span the mainland US and are consistently defined over time. The analysis focuses on 37 occupations, the respective 1950 codes are: Treatment group: 1, 5, 31, 51, 57; High income 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; Workers in other leisure activity placebo group: 4, 6, 77, 91, 732, 750, 754, 760, 784.

The variables used in the main analysis are: incwage, occ1950 (in combination with empstat), wkswork2, hrswork2. The wage data refers to wages in the previous calendar year. This data is first available in the 1940 Census. And in 1950 the income questions are only filled in by a subset of "sample-line" individuals. The IPUMS extracts are mostly sampled from these sample-line individuals and hence wage data is largely available. I convert the wage variables to real 1950 USD. The top-code bites above the 99th percentile of the US wage distribution in all years and we can therefore compute the share of workers in the top percentile.

Control variables are: median age & income, % female, % minority, population density, and trends for urban areas. Most variables are available consistently throughout the sample period. Income and education are only available from 1940 onwards. The Census race question includes changing categories and varying treatment of mixed-race individuals. I use the IPUMS harmonized race variable that aims to correct for those fluctuations. Additionally, I compute I compute the share of entertainers who move for each labor market. Note that the definition of mobility varies across Census vintages. Moreover, it does not distinguish between moves within and across labor markets. IPUMS aims to harmonize differences across Census vintages, and I use their harmonized variable. While such a measure is noisy, classic measurement error will not bias the results but rather inflate standard errors, as we use the variable as an outcome variable.

Employment Number of workers are based on labforce and empstat. Both variables are consistently available for those aged 16 years and older. Hence the sample is restricted to that age group. Occupation is recorded for ages older than 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 that 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 people do not report an occupation and thus do not fall into the sample of interest.[24] Second, the rate of unemployment is modest

[24]The unemployed may report an occupation if they have previously worked. I construct an alternative employment series that includes such workers for the entire sample period.

compared to that of employment and thus including the unemployed does not dramatically change the numbers.

I use the IPUMS 1950 occupation classification (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 as follows: The pre-1950 samples use an occupation system that IPUMS judges to be almost equivalent. For those samples IPUMS states that as: "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 and occupations IPUMS therefore relies on the raw data. There are, however, a few changes that do affect the occupation classifications:

- *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), all other entertainment professions are unaffected. Among the placebo occupations, a few new occupations categories are introduced in 1950.

- Changes *for the 1960–1970 period:* 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. The athletes category is discontinued in 1970 and the analysis therefore only uses this occupation until 1960. For the "Entertainers nec" category roughly 9,000 workers that were

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.

> previously categorized as “professional technical and kindred workers” are added along with a few workers from other categories in 1970. These added workers account for roughly 40% of the new occupation group. The occupation-specific year effect ought to absorb this change. I have performed additional robustness checks excluding 1970 or occupation groups and find similar results and the results are robust to this. Among placebo occupations, the “floor men” category is discontinued in 1970.

The industry classification also changes over time. The analysis uses 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 from 1960 to 1970, it falls back to around 90,000.

Pareto Approximations In some of the robustness tests I use Pareto extrapolations for top incomes beyond the top code. This follows a large literature that uses such approximations to measure the top tail of the income distribution (e.g., Kuznets and Jenks 1953; Atkinson, Piketty, and Saez 2011; Atkinson and Piketty 2010; Blanchet, Fournier, and Piketty 2017; Piketty and Saez 2003; Feenberg and Poterba 1993). If wages are Pareto distributed the distribution is pinned down by two parameters, the “Pareto coefficient” and the scale parameter. The cumulative distribution function of a Pareto distribution is: $1 - F(w) = (w/\omega)^{-1/\alpha}$, which is linear in logs. And the expected income for a person with top-coded income $\bar{y}$ is $E(y) = \frac{\alpha}{\alpha-1}\bar{y}$. For a top-coded observation, we can thus compute the expected income: it is k times the top-code and k is pinned down by the Pareto coefficient of the income distribution. The shape parameter conventionally used for the US income distribution is around $\alpha = 3$ and hence $k = 1.5$ (see e.g., Juhn, Murphy, and Pierce 1993; Lemieux 2006; Autor, Katz, and Kearney 2008).

An alternative approach is to estimate the α coefficient in the relevant data. Such coefficients can be calculated in a relatively straight-forward manner, since the wage distribution is log linear, the slope and intercept of this line capture the two key parameters of the distribution (α, ω). In principle, only two data points are enough data to recover the slope and intercept of the Pareto distribution. In practice, however, such estimates are extremely noisy and to improve the precision of the estimation, I restrict the sample to locations with at least 20 entertainers. The Pareto coefficient is given by $\alpha_{i,j} = [ln(income_i) - ln(income_j)] \,/\, [ln(rank_i) - ln(rank_j)]$. Using observations below the top code, I compute these Pareto coefficients for each local labor market and year and then impute unobserved incomes between observations from the estimated income distribution. With this approach I obtain the full entertainer wage distribution for each local labor market and year. I then use the data to calculate local top income shares, making use of the the fact that top income shares of a Pareto distribution are given by $S_{p\%} = (1-p)^{\frac{\alpha-1}{\alpha}}$.

B.2 Summary Statistics

Table B1 reports summary statistics for the baseline local labor market sample. This covers the 722 local labor markets for four Censuses (1940-1970), and thus 2,888 observations. The first set of results report statistics on the availability of television. The table reports averages for the full sample period. Since local filming only took place for a relatively short time period, the variable is zero in most years and the average number of TV stations is 0.02. At the time of local filming in 1949, filming occurred in around 5% of local labor markets through on average 1.78 stations. TV signal covers 60% of locations on average and signal coverage expands from no signal in 1939 to full coverage in 1969. The suitability of a location for filming is summarized by "local filming cost," and the data show the strong pull to concentrate filming when location decisions are unconstrained. The proxy for local comparative advantage is the number of movie productions in this local labor market in 1920. Most places had no movie sets, and only 16 locations produced at least 1 movie, with only LA producing more than 20 films. The average audience entertainers could attain was 72 million individuals. This is however skewed by the huge audiences in the national TV era. Before national TV, the average market reach is 62,000 individuals in 1949, while theater capacity of the pre-TV era only ranges from 400 to 12,000 individuals. Data on theater capacity is missing for 116 local labor markets, 16% of the sample.

Table B1: Summary Statistics

	No. of observations	Mean	S.D.
		Television	
Local TV stations	2,888	0.02	0.25
Local filming cost	2,888	0.14	1.36
Show audience (1,000s)	2,656	72,811	66,719
Show revenue ($1,000)	2,656	4,182,516	3,834,174
TV signal (%)	2,888	60	0.49
		Entertainment	
Employment in leisure activities	2,888	2,468	8,540
Employment in performance entertainment	2,888	177	936
Wage 99th percentile of entertainers ($)	1,435	5,704	4,576
Fair visits (thsd.)	8,664	25	109
Fair ticket receipts ($1,000)	8,664	2.94	1.89
Grandstand show receipts ($1,000)	8,664	1.64	0.97
Rides & carnival receipts ($1,000)	8,664	0.92	7.50
		Demographics	
People (1,000)	2,888	229	658
Workers (1,000)	2,888	86	264
Median income ($)	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 (CZs) over four decades. The 99th wage percentile is only computed for the larger local labor markets, see the text for details. The data is decadal, except *Fair* data, which is annual from 1946 to 1957. *Show audience* and *Show revenue* refers to the largest shows feasible in a CZ (see text for details), and no data are available for some CZs. *Median income* is missing in one CZ in 1940. *Urban Share* and *Filming Cost* are held fixed throughout the sample. Source: US Census 1940–1970, *Billboard* magazine 1946–1956.

Turning to entertainers, the average local labor market employs 177 performance entertainers during the sample period but there is again considerable heterogeneity across local labor markets (see demographics). Most important in the analysis are the local labor markets where TV filming took place, which have on average a little over 2,000 performance entertainers. Employment in all other leisure-related activities (i.e., including in bars and restaurants and in interactive leisure activities) is about 2,500 individuals in an average local labor market. The 99th percentile of the entertainer wage distribution averages close to \$5,700. Data on county fairs reports average attendance and spending in three categories: entrance tickets, shows, and rides and carnival purchases (e.g., candy, popcorn). These data show that county fairs are a popular event, with the average fair attracting about 25,000 visitors. These data are 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.

B.3 Robustness Tests

B.3.1 Monopsony Effects

A closely connected issue to the rise of market scale is the simultaneous rise in superstar firms and monopsony power. In many modern contexts SRTC may be associated with rising monopsony power, since a small number of technology companies control access to such technologies. The entertainment setting offers a unique setting to test this interaction of superstar effects and monopsony power. Government entry restrictions generate quasi-experimental variation in the number of competing local TV stations and thus allow me to identify the impact of labor market competition. First, as a benchmark I estimate the DiD in 2 with a dummy for TV filming as treat-

Table B2: *Effect of Competition in Local Labor Markets*

	(1)	(2)	(3)
	Entertainer among Top 1% of US Earners		
Local TV station (dummy)	5.90	0.75	-0.57
	(3.06)	(1.91)	(0.36)
Multiple local TV stations (dummy)		9.07	10.37
		(4.99)	(4.70)
Blocked competitor (dummy)			1.43
			(2.10)
Samples	full	full	full
No. of CZ cluster	722	722	722
Year–Occupation FE	Yes	Yes	Yes
CZ FE	Yes	Yes	Yes

[Note] The table shows the effect of competition between local TV stations. The regressors are a dummy with value one, respectively if a location has a TV station (Local TV station), a location has multiple TV stations (Multiple local TV stations) and a location has the entry of a second station blocked by the rollout interruption (Blocked competitor). For other specification details and sources see Table B5, Panel B.

ment variable. This captures the average effect of television stations and we do see again substantial growth in highly paid entertainers (Table B2, column 1). Next, we distinguish between places with a single monopsony TV station and places with multiple stations. The results show a marked difference between monopsonistic and competitive labor markets. Markets with a single TV station see almost no top income growth, while in markets with competing TV stations top incomes increase sharply. These results also hold when I narrow in on the variation from the rollout interruption experiment. Places where the entry of competing stations is blocked continue to look like monopsony locations (Table B2). These findings emphasize the importance of competition for superstar effects. The growing market scale only translates into rising top pay if employers are competing for talent.

B.3.2 Wage Quantile Effects

This section quantifies the impact of TV on wage percentiles at the top of the entertainer distribution. To do so, I compute the top percentile of local entertainer wages. In most cases, this approach uses the highest observed entertainer wage in the local labor market as proxy for the top percentile. I restrict the sample to larger labor markets to limit noise in this measure.[25] Specifically, I use the "rollout interruption sample" from above and thus compare places where television was launched to ones where launches were blocked during the rollout interruption. The results are robust to alternative sample choices. I then repeat the DiD analysis of equation 2, using the log of these wages as the outcome variable.[26] In a deviation from equation 2, these regressions are run at the CZ-year level. This change is necessary because quantiles are not additively separable into sub-groups and we therefore cannot disaggregate wage quantiles by occupations.[27]

I find a sharp and sizable increase in top entertainer incomes with the launch of a local TV station. Panel A in Table B3 shows an increase in the 99th percentile by 18 log points, or approximately 20%. A 20% wage increase is large in any context, but it is a particularly striking increase given that the regression includes year fixed effects and the results are thus on top of average wage growth. The 95 percent confidence interval ranges from a 5% growth to 35% and is thus relatively large. Allowing for broader samples that introduce additional control areas increases the precision and yields similar point estimates.

In 10% of cases the 99th percentile wage exceeds the top code, and I show that results are robust to using alternative methods from the literature

[25] If over 100 individuals are sampled, I use the sample weights to compute wages at the 99th percentile (19% of observations).

[26] This approach amounts to a quantile DiD estimate (Chetverikov, Larsen, and Palmer 2016).

[27] Note that the aggregated regressions use fewer observations but have the same power as disaggregated specifications as the number of CZ clusters stays the same.

to adjust for top-coding. The first set of specifications in Panel A make no adjustments for the top code and thus ignore earnings growth beyond the top-code level. This will underestimate the true top earning growth and, as a result, likely provides a conservative estimate for the magnitude of superstar effects. In Panel B I use the fixed-multiple approach to top-coding and assume a constant multiplier of 1.5 (see e.g., Juhn, Murphy, and Pierce 1993; Lemieux 2006; Autor, Katz, and Kearney 2008). In Panel C I use local Pareto approximations to impute the top coded wages.[28] As expected, imputing incomes beyond the top code raises the magnitude of the effects somewhat. The estimates remain in the same ballpark; at the 99th percentile income growth is 20% to 30%. Specifications that add controls for demographics or location specific trends yield similar results.

Table B4 shows the impact on the income shares of top entertainers.[29] To compute such top income shares, we need information on the full population or a parametric assumption about the shape of the top income tail. In line with the wider literature on top incomes shares and Table B3, I use Pareto approximations to compute such shares.[30] Such imputations are less reliable in small samples and the regressions use weights that put more weight on larger CZs. Additionally, I test whether the results are robust to alternative sample restrictions that exclude small CZs. Columns 1-3 compute top income shares in all cells with at least 20 entertainers and Columns 4-6 use the "rollout interruption sample," focusing on areas with local television filming or affected by the interruption.

The launch of a TV station increased the top 1% income share by 45 log points, or 57% (Table B4, column 2). In line with Proposition (b)—which suggests that the growth in these shares escalates toward the top of

[28]For details on the procedures, see Online Appendix B.1.

[29]Top income shares are widely used to measure inequality at the top. See, for example, Piketty and Saez 2003; Piketty 2014.

[30]Table B3 uses Pareto approximations for top-coded observations only, here we additionally require such approximations in all cells without information on the full population.

Table B3: *Effects on the 99th Percentile*

	(1)	(2)	(3)
	Ln(99^{th} *Percentile of Entertainer Wages)*		
	Panel A: No Imputation		
Local TV stations	0.182	0.189	0.147
	(0.078)	(0.079)	(0.1168)
	Panel B: Fixed Multiple Imputation		
Local TV stations	0.213	0.218	0.171
	(0.085)	(0.087)	(0.1235)
	Panel C: Pareto Imputation		
Local TV stations	0.283	0.277	0.237
	(0.095)	(0.089)	(0.131)
Sample	interrupted	interrupted	interrupted
No. of CZ clusters	112	112	112
Year FE	Yes	Yes	Yes
CZ FE	Yes	Yes	Yes
Demographics	–	Yes	–
Local labor market trends	–	–	Yes

[Note] The Table tests the effect of local TV launches on entertainer top incomes, using the quantile DiD estimator developed by Chetverikov, Larsen, and Palmer (2016). Outcome: ln(99th percentile of local entertainer wages) computed at the CZ-year level. The panels differ in how they adjust for top-coding: Panel A makes no adjustments, Panel B uses the fixed multiple approach and multiplies top-coded observations by 1.5, Panel C uses local Pareto approximations. The control variables are as in Table B5. The sample uses the "Rollout Interruption sample" of Table B5 Panel C and covers 112 CZ cluster over 4 years and 400 CZ-year observations. 48 observations are missing due to cell size restrictions in computing Pareto imputations. Regression run at the CZ-year level since the 99th percentile cannot be disaggregated by occupation. Observations are weighted by local labor market population. Standard errors are reported in brackets and are clustered at the local labor market level. Source: US Census 1940-1970.

Table B4: *Effect of TV on Top Income Shares in Entertainment*

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln(Share of Income)			*Ln(Share of Income)*		
	top 0.1%	top 1%	top 10%	top 0.1%	top 1%	top 10%
Local TV stations	0.68	0.45	0.23	0.47	0.32	0.16
	(0.19)	(0.12)	(0.06)	(0.20)	(0.14)	(0.07)
P-value:						
$\Delta\ y = \Delta$ *top 1%*	0.02	—	0.00	0.24	—	0.00
Sample	big CZs	big CZs	big CZs	interrup.	interrup.	interrup.
No. of CZ cluster	346	346	346	112	112	112
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
CZ FE	Yes	Yes	Yes	Yes	Yes	Yes

[Note] The table shows the effect of local TV stations on top income shares in entertainment. Outcomes: The top p% is the share of income going to the top p percent of entertainers in a given local labor market–year. Estimates are based on a DiD specification across CZ-year cells. Top income shares are calculated using local Pareto approximations. Column 1 to 3 does these interpolations in all CZ-year cells with at least 20 entertainers, which leads to a sample of 1,061 CZ-year observations and 346 CZ cluster, while columns 3 to 6 show results for the smaller "rollout interruption sample," as in Table B3 Panel C. *P-value:* test if change in outcome variable is the same as the change in top 1% income shares. The test is implemented in a regression with the ratio of top income shares as outcome variable. Observations are weighted by cell-size. Standard errors are clustered at the local labor market (CZ) level. Sources: US Census 1940–1970.

the distribution—I find that income gains for the top 1% are substantially larger than among the broader top 10% (for which income share increases by 23 log points) but are smaller than for the top 0.1% (an increase of 68 log points). These results can be used to formally test Proposition (b); a test for equality of growth rates is strongly rejected. Similar results hold in the more-restricted "rollout interruption sample" in columns 4-6.

B.3.3 Alternative Normalization

The baseline analysis studies the share of entertainers in the top 1%. The denominator of the share is fixed at the average employment in the labor market to prevent spurious effects from exit. Table B5 shows the results without this normalization. Panel A reports the baseline results, Panel B are per Capita counts and Panel C raw counts.

B.3.4 Tripple Difference

We can combine placebo and entertainment occupations to run a triple difference analysis. In a first step I pool placebo and entertainment occupations and allow a TV station launch to have different effects on the two groups. Results show that only entertainers benefit from the TV launch (Table B6, Column 1). The estimated effect on performance entertainers remains similar to the baseline DiD regression. Column 2 allows for a separate impact of television for each occupation of the placebo occupations, which shows that entertainers are indeed different from all other placebo occupations. Finally, I run the full triple difference regression. In this regression, the treatment varies at the time, labor market, and occupation level, which allows me to control for pairwise interactions of time, market, and occupation fixed effects and thus capture local demand shocks that happen to coincide with TV launches. 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. Such shocks could lead to an upward bias in the estimates of a DiD set up but will now be captured by the location-specific time effects.

Table B6: Earning Effect of TV Launch—Triple Difference Analysis

	Share in Top 1%		
	(1)	(2)	(3)
Local TV station × Placebo occupation	-0.41		
	(0.47)		
Local TV station × Performance entertainer	4.87	4.87	4.17
	(2.16)	(2.16)	(1.57)
Local TV station × Interactive leisure		-3.40	
		(1.29)	
Local TV station × Bars & restaurants		-3.80	
		(1.84)	
Local TV station × Professional services		5.23	
		(4.86)	
Local TV station × Medics		-3.24	
		(1.52)	
Local TV station × Engineer		-1.12	
		(1.23)	
Local TV station × Manager		3.55	
		(2.21)	
Year–Occupation & CZ FE	Yes	Yes	–
Pairwise interaction: Location, year, occupation FE	–	–	Yes

[Notes] The table shows triple difference results of local TV stations on top earners. Data and specification are as in B5. The number of CZ–occupation–year observations is 100,308.

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 confounders. The introduction of a "superstar technology" thus has a large causal effect on top incomes, and this effect is unique to the treated group.

B.3.5 Pre-Trend

A challenge for estimating pre-trends with this sample is that wage data in the Census is first collected in 1940. 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 IRS tax return data. In 1916 the IRS published aggregate information on top earners by occupation-state bins, including data for actors and athletes. I link the Census data with the tax data and run the regressions at the state level. Table B7 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 the treatment and control groups for the various years. The results show a clear spike in 1950, the year of local television filming. Looking at pre-trends, there is no significant pre-trend, in part because the standard errors are large. If anything, the treated areas seem to be on a slight relative downward trend in the pre-period, in line with the well known aggregate decline of top incomes during the 1930s. Even if we take this insignificant trend at face value, the pre-trends could go in the opposite direction and cannot explain the identified positive effect of TV launches.

B.3.6 Leapfrogging and the Talent Distribution

A stable distribution of talent implies that star entertainers remain at the same income rank and entertainers who appear at the top of the distribution after television should also appear at the top before television. Instead, we would expect leapfrogging in the distribution if stations relied on a different type of talent.

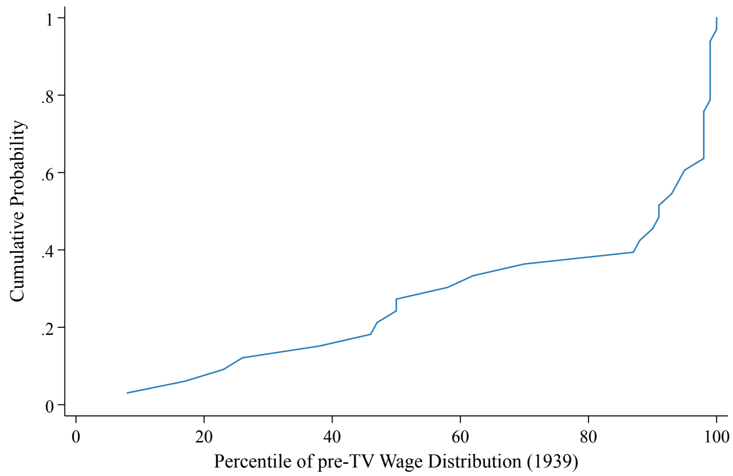
I build a small panel on the work history of TV superstars to study leapfrogging. The data on TV stars comes from the 1949 "Radio and Television Yearbook" which publishes an annual "Who is Who" in television—a list of stars similar to modern Forbes lists. The data covers the top 100 or so most successful TV entertainers and their demographic information (e.g., names, TV station employer, birthdays and place of birth) but not income. To obtain information on their pre-TV careers, I link this data to de-anonymized records of the 1940 Census. This link is based on names and additional demographic information (e.g., place of birth, birth year, parental information) and I can uniquely identify 59 of these TV superstars in the Census.[31] While the data is inevitably imperfect, it offers a rare window into the background of the stars of a profession and allows me to study the background of the group that benefitted most from the SRTC of television.

The panel data shows no leapfrogging and instead reveals that television stars where already disproportionally high-paid before television (see Figure A4).[32]

[31]To maximize the match rates, I additional hand-collected biographic information on the TV stars from internet searches. Demographic information on entertainer stars is unusually well documented due to the large amount of fan interest. As a result, I achieve a 70% unique match rate among the 68 records with birth-year information, while a few cases are matched without birth-year information.

[32]The data includes workers at different stages of the life-cycle. To avoid that such factors distort the wage rank of younger workers, I compute the wage ranking after residualizing wages for age, education and gender. This roughly corresponds to ranking individuals within their peer group.

Figure A4: *Wage Rank of Future TV Stars in the 1939 US Wage Distribution*



[Note] The Figure shows the wage ranks of TV stars before they became TV stars. TV stars are defined in the 1949 "Radio and Television Yearbook" 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: See Text.

References

Atkinson, Anthony B and Thomas Piketty (2010). *Top Incomes: A Global Perspective*. Oxford University Press.

Atkinson, Anthony B, Thomas Piketty, and Emmanuel Saez (2011). "Top Incomes in the Long Run of History." In: *Journal of Economic Literature* 49.1, pp. 3–71.

Autor, David H. and David Dorn (2013). "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market." In: *American Economic Review* 103.5, pp. 1553–97.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney (2008). "Trends in U.S. Wage Inequality: Revising the Revisionists." In: *Review of Economics and Statistics* 90.May, pp. 300–323.

Blanchet, Thomas, Juliette Fournier, and Thomas Piketty (2017). "Generalized Pareto Curves: Theory and Applications." In: *Working Paper* March.

Chetverikov, Denis, Bradley Larsen, and Christopher Palmer (2016). "IV Quantile Regression for Group-Level Treatments, With an Application to the Distributional Effects of Trade." In: *Econometrica* 84.2, pp. 809–833.

Feenberg, Daniel R. and James M. Poterba (1993). *Income Inequality and the Incomes of Very High-Income Taxpayers: Evidence from Tax Returns*. Vol. 7. January, pp. 145–177.

Fenton, George and Felix Koenig (2020). "Labor Supply and Innovation in Entertainment: Evidence from TV." In: *Working paper*.

Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce (1993). "Wage inequality and the rise in returns to skill." In: *Journal of Political Economy* 101.3, pp. 410–442.

Kuznets, Simon and Elizabeth Jenks (1953). *Shares of Upper Income Groups in Savings*. Vol. I, pp. 171–218.

Lemieux, Thomas (2006). “Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?” In: *American Economic Review* 96.3, pp. 461–498.

Manson, Steven, Jonathan Schroeder, David Van Riper, and Steven Ruggles (2017). *IPUMS NHGIS: Version 12.0.* Tech. rep. Minneapolis: University of Minnesota.

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

Piketty, Thomas and Emmanuel Saez (2003). “Income Inequality in the United States, 1913-1998.” In: *The Quarterly Journal of Economics* 68.February, pp. 1–39.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek (2017). *Integrated Public Use Microdata Series: Version 7.0.* Tech. rep. Minneapolis: University of Minnesota.

Table B5: *Effect of TV on Top Earning Entertainers*

	(1)	(2)	(3)
	Panel A: Entertainer among Top 1% of US Earners (% of Entertainers)		
Local TV stations	4.14	4.31	5.93
	(1.26)	(1.27)	(2.21)
Increase on baseline	92%	96%	132%
	Panel B: Entertainer among Top 1% of US Earners (Per Capita in 100,000s)		
Local TV stations	0.40	0.40	0.31
	(0.10)	(0.10)	(0.10)
Increase on baseline	133%	133%	103%
	Panel C: Entertainer in US top 1% (Raw Counts)		
Local TV stations	30.91	32.09	19.31
	(8.92)	(9.92)	(8.31)
Increase on baseline	199%	207%	124%
Sample	full	full	full
No. of CZ cluster	722	722	722
Year-Occupation FE	Yes	Yes	Yes
CZ FE	Yes	Yes	Yes
Demographics	–	Yes	–
CZ level trends	–	–	Yes

[Note] The table shows DiD results from estimating equation 2, regressing the respective outcome variables on the number of TV stations in the local area, each cell is a separate regression. Outcomes: Panel A, share of local entertainers among the top 1% of the US income distribution; Panel B, local entertainers among the top 1% of the US income distribution per capita in 100,000s; Panel C, raw counts local entertainers among the top 1% of the US income distribution. All regressions control for commuting zone (CZ), occupation specific time fixed effects and local filming cost in years after the invention of the videotape. Entertainers are actors, athletes, dancers, entertainers not elsewhere classified, musicians. Column 2 controls for median age & income, % female, % minority, population density, and trends for urban areas. Column 3 controls for a separate linear trend for each CZ. Sample: include 13,718 observations in 722 CZs, 5 occupations over four years, except for the athlete occupation, which is available for three years. Demographic data is missing for one CZ in 1940 and thus reduces the sample in column 2. *Increase on 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 B7: Effect of TV on Top Earning Entertainers—State Level

	Share in Top 1%		
	(1)	(2)	(3)
Local TV station ×(1916)			8.31
			(5.97)
Local TV station × (1940)			0
			-
Local TV station × (1950)	20.94	20.18	23.32
	(8.09)	(7.36)	(7.27)
Local TV station × (1960)			1.70
			(2.60)
Local TV station × (1970)			8.90
			(2.95)
Years	1940–1970	1916–1970	1916–1970
Year & State FE	Yes	Yes	Yes
No. of observations	912	1008	1008

[Notes] The Table shows results of pre-trend tests. Data and specification are as in B5, Panel A except that the data is now aggregated at the state-year-occupation level. Standard errors are clustered at the state level and appear in parentheses. Each row represents a separate DiD regression. Column 1 estimates the baseline specification of Table B5 in the aggregated data, column 2 extends the time period and column 3 introduces leads and lags of the treatment. 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 1940. Source: US Census (1940–1970) and IRS in 1916.