ONLINE APPENDIX

A1. TECHNICAL CHANGE AND SUPERSTAR EFFECTS: EVIDENCE FROM THE ROLLOUT OF TELEVISION

A2. by Felix Koenig

APPENDIX: Derivations

B1. 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,

$$Y(s_i, t) - w(t) \geq Y(s_i, t') - w(t') \quad \forall \ t' \epsilon [\underline{t}, \overline{t}]. \tag{B1}$$

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 B1 by ϵ and let $\epsilon \to 0$.

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

$$\frac{\partial w}{\partial t} = \frac{\partial Y(S_i, t)}{\partial t}. \tag{B2}$$

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

$$Y(s_i, t) - w(p) \geq \psi^{res} \quad \forall \ p\epsilon[\bar{p}, 1] \tag{B3}$$

$$w(p) \geq w^{res} \quad \forall \ p\epsilon[\bar{p}, 1]. \tag{B4}$$

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 (B2) as

$$\frac{\partial w}{\partial t} = \frac{\pi}{\phi} s^{\frac{1}{\phi}} t^{\frac{1}{\phi}-1} = \frac{\pi}{\phi} t^{\frac{1}{\xi}-1}, \tag{B5}$$

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:

$$w(t) = \int_{\underline{t}}^{t} \frac{\partial w}{\partial t} = \frac{\pi\beta}{\alpha+\beta} t^{-1/(\xi\beta)} = \frac{\pi\beta}{\alpha+\beta}[p_w]^{-1/\xi}. \tag{B6}$$

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

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

B2. *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 $\varepsilon_D > \frac{\partial M/\partial\xi}{(\alpha+\beta)\kappa}$).[16] 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).

[16]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 $[w^{-\xi} - w'^{-\xi}] + [w^{-\xi}(ln(w) - 1) - w'^{-\xi}(ln(w') - 1)] > [w^{-\xi} - w'^{-\xi}]$.

APPENDIX: Empirics

C1. 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.[17] 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 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.[18] The share in market m at time t is:

$$p^{\omega^{99}}_{m,t} = \frac{\sum_{i \epsilon I} E_{i,m,t}}{\overline{E_t}}, \tag{C1}$$

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%.[19] 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

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

[18] 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.

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

labor market counts.[20] 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 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.

[20] 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 $p^{\omega^{99}}_{m,t}$ 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 C.C3).

Worker data

Data is provided by the Integrated Public Use Microdata Files (IPUMS, Ruggles et al. 2017; Ruggles et al. 2021) of the US decennial Census from 1920 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. The core of the analysis therefore focuses on the 1930-1970 period. 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.[21] Second, the rate of unemployment is modest 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 previously categorized as

[21]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. 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.

> "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 $S_{p\%} = (1-p)^{\frac{\alpha-1}{\alpha}}$.

C2. *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.

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. 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.

C3. *Robustness Tests*

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

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
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 ($)	1,435	5,704	4,576
	Demographics		
People (1,000)	2,888	229	658
Workers (1,000)	2,888	86	264
Median income ($)	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. *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.

a dummy for TV filming as treatment 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.

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	-0.55
	(3.06)	(1.91)	(0.32)
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.

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.[22] 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.[23] 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

[22]If over 100 individuals are sampled, I use the sample weights to compute wages at the 99th percentile (19% of observations).
[23]This approach amounts to a quantile DiD estimate (Chetverikov, Larsen, and Palmer, 2016).

therefore cannot disaggregate wage quantiles by occupations.[24]

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 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.[25] 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.[26] 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.[27] 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 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

[24] 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.

[25] For details on the procedures, see Online Appendix C.C1.

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

[27] 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.

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.

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.

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 **??**, 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.

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.

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.

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.

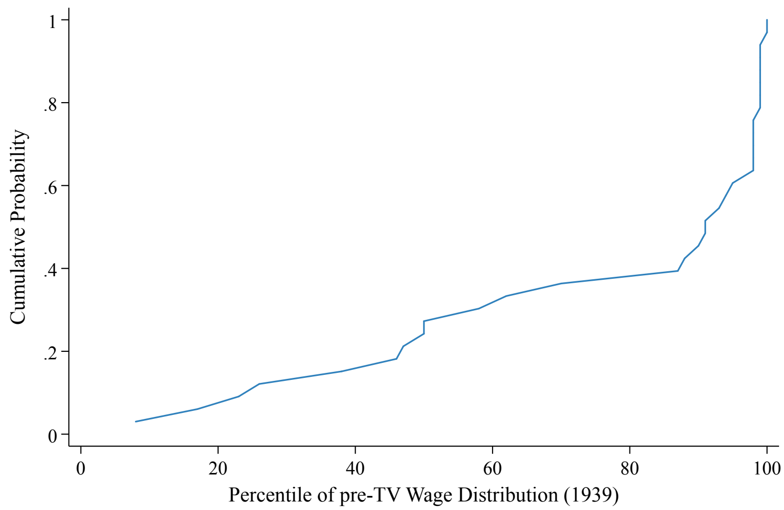
I build a small panel on the work history of TV superstars to study leapfrogging. The data on TV stars comes from the 1949 "Radio and Television Yearbook" which publishes an annual "Who is Who" in television—a list of stars similar to modern Forbes lists. The data covers the top 100 or so most successful 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.[28] 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 A1).[29]

[28] 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.

[29] 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 A1. : 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, Catherine A. Fitch, Ronald Goeken, J. David Hacker, Matt A. Nelson, Evan Roberts, Megan Schouweiler, and Matthew Sobek (2021). *IPUMS Ancestry Full Count Data: Version 3.0 [dataset]*. Tech. rep. Minneapolis, MN: IPUMS.

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 B3—: Effects on the 99th Percentile

	(1)	(2)	(3)
	Ln(99th Percentile of Entertainer Wages)		
	Panel A: No Imputation		
Local TV stations	0.182	0.189	0.147
	(0.078)	(0.079)	(0.118)
	Panel B: Fixed Multiple Imputation		
Local TV stations	0.213	0.218	0.171
	(0.085)	(0.087)	(0.125)
	Panel C: Pareto Imputation		
Local TV stations	0.283	0.277	0.237
	(0.095)	(0.089)	(0.132)
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.

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	5.92
	(1.26)	(1.27)	(2.20)
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	0.30
	(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 B6—: Earning Effect of TV Launch—Tripple 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

Note:
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.

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

Note: 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.