Superstar Earners and Market Size: Evidence from the Entertainment Industry*

Felix Koenig †

April 28, 2017

Abstract

"Superstar theory" argues that growth in top incomes is driven by a growth in market size. Despite the prominence of this superstar idea, there is no causal evidence for the effect of market size on top incomes. This paper makes use of a historical quasi-experiment to provide a causal test for "superstar effects". The roll-out of television in the middle of the 20th century increases the size of the market an entertainer can serve. Signal interference of broadcasting signal gives rise to quasi-random variation in the location of television production. This allows me to estimate the causal effect of a market size enhancing technology on the wage distribution. I build a novel dataset that digitizes detailed entertainment industry records, information on the location of television and the entertainment market size. I link this to employment and wage records from the US census. The introduction of TV has a large and significant effects on the distribution of incomes. The number of entertainers in the top 1% of the income distribution doubles, while the market size increases more than 4-fold. The causal effect of market size on top earners is economically large and significant. At the same time the causal estimate are smaller than the OLS estimates. The estimates thus suggest that the correlation of top income and market size over states the role of superstar effects.

*I am grateful to many people in particular my supervisor Steve Pischke.

†London School of Economics, Centre for Economic Performance, 32 Lincoln's Inn Field, WC2A 3PH London, UK.

Contents

1 Introduction

Top incomes have grown rapidly in many countries. Surprisingly, there is little causal evidence on what drives this rise in top income inequality. Rosen famously proposed a "superstar theory" that features large incomes differences in the top tail of the distribution. The idea of the superstar effect is that talent becomes more valuable in larger markets. A good idea, for instance, is more valuable if used on a large scale. On the contrary mistakes are particularly costly if they affect large operations. The value of talent thus rises in large markets. The growth in available market size may therefore plausibly increase the return to talent. According to Kaplan and Rauh this market size effect is the main driver behind the observed rapid growth in US top incomes in recent decades Kaplan and Rauh (2013). Market size has risen sharply in recent years. Among other factors mass-production technology, communication technology and falling trade barriers have contributed to this trend. Industry champions have emerged in many sectors and markets have moved closer to winner takes it all markets. This growth in market size has coincided with the increase in top incomes. However, no causal link has so far been established. The aim of this paper is to provide a causal test for the effect of market size on income at the top of the distribution.

The empirical analysis of superstar effects has been mainly descriptive. Krueger (2005) documents fast growth of top incomes among rock-stars. He concludes that the greater income concentration at the top mimics the change in the overall income distribution. There is a large literature that documents a positive correlation between CEO pay and market size. Most recently, two papers build on Rosen's superstar model to derive empirically testable predictions. They link firm size to CEO pay. Both show that CEO pay is significantly positively correlated with market capitalisation of the employing firm. Gabaix and Landier (2008) look at the correlation of firm value and CEO pay in the cross-section and time series and find a robust positive relation. Terviö (2008) compares CEO wages in industries of different size. Firms that produce a product with a larger product niche tend to have higher CEO wage.[1] Both papers stand in a long tradition of papers that find a strong correlation between firm value and CEO pay (for an overview of this literature on CEO pay see Murphy (1999)). A related literature looks at scale effects *within* firms. Communication technology changed the way management and supervision can be done and arguably made it cheaper to manage larger teams. This in turn changes the team size and increases the return to talent for managers. As a result earning dispersion within a firm increase (e.g. Garicano (2000)).

[1]Differences in wages across product niches could come from a variety of sources such as differences in employee characteristics, barriers to entry, firm specific wage arrangements, differences in skill requirement and other factors that determine the marginal revenue product or wage of a worker.

To the best of my knowledge the literature has not provided any causal evidence for the effect of market size on top incomes. Estimating the causal effect of market size is complicated by a number of factors. For one the market size tends to be endogenously determined. Managerial decisions, such as upgrading of firm equipment and investment in innovation play an important role in the expansion of market size. These decisions are related to unobserved characteristics of the individual and thus endogenously determined. The challenge is therefore to isolate growth in market size that is driven by exogenous factors and growth that is a result of endogenous firm decision. This paper will use quasi-experimental variation in the available market size to trace out the causal effect of growing markets for top incomes.

A second challenge is to measure market size. This is particularly relevant to link the estimates to a model of "superstar effects." The model distinguishes between observed scale of operation, such as profits, wages or revenue, and a primitive that is independent of worker talent, such as market's potential or capacity. The latter drives the superstar effect. In the context of entertainment this capacity concept is well defined. Each performance venue has a fixed capacity that an entertainer can potentially fill. I digitize historical records of more than 3,000 theaters and performance venues to measure the potential audience of a performance. For later periods I make use of detailed technical and geospatial information to calculate how many people can be reached from each TV station. These records are matched with data on TV ownership. The availability of data on a well defined measure of market size is another advantage of an analysis of the entertainment industry.

This paper will use exogenous variation in the introduction of television as a market size shock in the entertainment sector. The size of TV audience exceeded the largest previously available venues size by factor 4 and thus provided a substantial market size shock. I will make use of problems with signal interference and government intervention that will lead to exogenous variation in the location of TV production. This will allow me to estimate the causal effect of market size for the wage distribution. To the best of my knowledge this paper is the first to estimate the causal effect of market size on top incomes. Television only spreads slowly across the US, halted by several regular interventions. Commercial television was launched in 1941. By 1950 only about 40% of US commuting zones have access to TV signal. In 1950 for instance people were regularly watching television in Sycharus, whereas the technology was not yet available in Denver. Only by the end of the decade was universal coverage achieved. The slow roll-out of TV thus gives rise to staggered introduction of a "superstar technology" across several markets. The analysis will make use of this variation in a difference in differences analysis. A second layer of exogenous variation is introduced by government regulation. This variation arises due to problems with the signal transmission. Television struggled with the shortage of frequency spectra from the start. The technology of the time could only receive and process signals in a relatively small range of frequency spectra. The government thus imposed heavy regulations around the location and equipment of TV broadcast aerials. For the years 1948-1952 the government even introduced a blanket ban on new aerials. This gives rise to externally imposed differences in the timing of treatment across markets. I digitize new data on the application process for TV broadcasting licenses. This shows a complete shut down in approvals in the late 1940s. A number of areas don't receive television due to the government license "freeze". The sharp discontinuity in approvals allows me to further strengthen my identification strategy. A third layer of exogenous variation is that television works for certain entertainment tasks but not for others. This task specific effect of the television allows me to run placebo tests. Television transformed the delivery of performance entertainment but didn't affect the delivery of a number of other leisure activities. Activities that are unaffected include for example bowling, eating out or sport lessons. Comparing the effect of TV on the workers in these different entertainment occupations therefore allows me to differentiate the effect of greater scalability from other location specific changes in the entertainment sector. Finally, I run a triple difference estimation. This analyses compares the effect of TV on the two groups and simultaneously controls for location specific year effects, occupation specific year and location specific occupation effects in a non-parametric way. This exercise rules out a number of alternative explanations and confirms that TV has a large causal effect on performance entertainers.

Another strength of this setting is that it allows me to separately identify the effect of the technology on the market size and on competition. In labor markets with television production entertainers can reach a bigger audience. Superstar theory predicts that this leads to greater concentration of income at the top. On the other hand some labor markets will experience added competition from the expanding signal of television antennas, but they don't have access to television production. This may result in cannibalizing effects of technological progress. In trade settings expansion in market size allows both goods to be exported as well as competition from foreign imports. Hence market size and the number of competitors increases for both markets simultaneously. In the television case these factors can be disentangled. Many labor markets only experience competition from television shows while only a few get access to the technology that allows to produce shows. In other words there are markets that only experience greater competition but don't have access to the new technology to expand the production scale.

The remainder of the paper is organized as follows. Section 2 presents a model of superstar effects, section 3 describes the roll-out of television in the US, Section 4 gives an overview over the data sources and variables, Section 5 presents the estimation results, section 6 concludes.

2 The superstar model

The superstar model features workers with heterogeneous talent. Talent is distributed according to some distribution function with percentile $p = F(p)$.[2] The workers can work in markets of varying sizes. In the context of entertainment, the markets can be thought of as a stage. For simplicity I assume that at each point in time only one individual can appear on each stage. . The output of a worker depends on the audience size S and is given by $Y(S, p)$ with $Y_S > 0$, $Y_p > 0$, $Y_{pp} < 0$ and $Y_{Sp} > 0$ where subscripts indicate derivatives. The last condition ensures that a highly talented becomes more valuable the bigger the market. This condition will generate positive assortative matching in equilibrium. A stage manager (or later TV entrepreneur) has to pick the optimal person from the talent distribution to perform on his stage. The problem for a stage manager with a stage of size S_i is given by:

$$max_p Y(S_i, p) - w(p)$$

where $w(p)$ is the wage for an entertainer from the p percentile of the distribution. The optimality condition is given by:

$$w'(p) = Y_p(S_i, p) \tag{1}$$

This function pins down the difference in wages as you move along the talent distribution. If we let i represent the percentile of the stage size in the distribution of stage sizes, assortative matching ensures $i = p$. Integration than gives the wage distribution. Wages are pinned down up to a constant that represents the outside option of the least talented worker that participates $w(0)$.

$$w(p) = w(0) + \int_0^p Y_p(S_j, j)dj \tag{2}$$

The wage of a worker at percentile p depends on all worker-stage pairs below him, but is independent of anything that happens at higher percentiles. Productivity changes at the bottom end of the distribution will thus effect everyone. The logic for this is that bigger venues will pay a higher price per talent unit to attract the best talent. Each venue

[2] We assume that the CDF is differentiable.

cares only about distinguishing itself from the next worse employer and thus pays a mark-up on their wage. In this sense all employers look "downward" in the distribution. An increase in wages at the bottom thus has a domino effect and will push up all wages. An increase in the productivity at the top will however not affect wages further down in the distribution. A second feature of this wage setting equation is that growth in productivity at all levels, including the outside option, will shift up all wages equally and leave relative wages unchanged.

An equilibrium in this setting requires: an assignment of entertainers to markets $S_i = \theta(p)$, a wage schedule $w(p)$, all stage managers maximize profits and markets clear. The definition of the equilibrium together with the first order condition allow to show what generates skewness in earnings in this model. The second derivative of equilibrium wages is given by:

$$w''(p) = \frac{\partial Y_p(\theta(p), p)}{\partial p} = Y_{pp}(S_i, p) + Y_{pS}(S_i, p)\frac{\partial S_i}{\partial p}$$

The last term in this equation is positive by the definition of Y and by assortative matching. This term thus provides the force that allows for a convex earning distribution.

Next, I will turn to analyzing the effect of a change in the size of available markets. To keep the model tractable I will impose some structure on the production function. I will assume that output Y is separable in p and S_i and takes the form $Y = S_i^\gamma g(p)$ where $g()$ is an increasing function.[3] Since there is no natural unit for talent one could pick an arbitrary measure T and make output multiplicative in that measure of talent $T(p) = g(p)$. Since those units would however be meaningless, I will keep production as a function of the percentile of the hired talent. This can be contrasted to a market without superstar effects. A simple case is $Y = S_i^\gamma + g(p)$. Here we turned off the complementarity of market size and talent, $Y_{pS} = 0$ and thus we won't get assortative matching. Workers are still paid their marginal product and there is a single market clearing price for talent. An increase in market size doesn't affect the marginal product of a worker and thus leaves wages unaffected.

Gabaix and Landier show that the model can be simplified to an easily estimatable equation. They assume that the firm size distribution is Pareto with tail parameter α, production is given by $Y = S^\gamma p$ and the talent distribution follows a "regular" continuous distribution. This allows us to write the wage as:

$$w(p) = D_{i^*} S_{i^*}^{\alpha/\beta} S_i^{\gamma - \alpha/\beta}$$

Here i^* is a reference firm and D is a size independent constant. Market size affects wages

[3]This assumption is standard in the literature and also used in all major papers on superstar effects: Rosen, Lucas, Tervioe, Gabaix & Landier.

through two channels. For one, all firms benefit from a greater average market size, captured by i^*. Second, the size of the available market for a given individual has an effect over and above the average market size.

The above equation becomes a linear equation in log market size after taking logs with δ a time fixed effect that captures average market size growth.

$$ln(w_i) = \delta_t + [\gamma - \alpha/\beta] * ln(S_i) + \epsilon$$

A number of studies estimate the effect of market size on CEO pay. These studies find that the coefficient on log market size is around 1/3 in the cross-section and around 1 in the time-series Gabaix and Landier (2008).

3 Entertainment on TV

To estimate the effect of market size on top incomes we will need to measure the available market size and track exogenous variation in market size. In the entertainment sector television provided a big shock to the market size. A single television production reached an audience an order magnitude bigger than the capacity of any performance venue. I will return to the measurement of market size below. First I will focus on the location and roll-out of television productions. The television shock was not experienced equally across labor markets in the US. Production of television went through two major phases. The first phase was a phase of local TV productions. In the early years the location of production was closely linked to the location of the broadcasting antenna. Filming took place near the antenna and was transmitted live. Storing or transporting productions was unpopular because recordings were of poor quality and expensive. The implication of this was that much content at this time was locally produced. The construction of new television antennas thus led to a shift in the availability market size for local entertainers. The first part of the analysis will thus use the variation introduced by the process of subsequent installation of television antennas across the US. Figure 4 shows a US map of the location of television stations in 1949. The phase of local television ends with the invention of the Videocassette Recorder (VCR) in 1956. The VCR offered a cheap way of storing television productions and TV productions become a transportable good. As a consequence local studios lost in importance and production of television moved away from broadcasting antennas. In terms of identification the rise and fall of local television provides nice quasi-experimental variation. The construction of a television antenna first provides a positive shock to market size, but this effect is removed after the invention of the VCR. Local labor market thus experience first a positive and then a negative shock in the available market size. This variation allows me to estimate the effect of market size on income at the top.

3.1 TV availability

The location of television broadcasting antennas will play a crucial role in the empirical analysis. It will therefore be important to understand how the location of these antennas came about. Commercial Television Broadcasting was launched in the US on July 1, 1941.[4] In the pioneer days of television, broadcasting antennas were set up that could transmit programs to the surrounding area. In 1941 television was broadcast in a handful of cities throughout the US.

TV antennas were subsequently set up in other cities. Such broadcasting stations were set up by private investors with the approval of the Federal Communication Commission (FCC). The location of antennas was thus not at random places throughout the US. Instead, antennas were set up in markets selected by entrepreneurs. The key concern for entrepreneurs was to be able to generate large advertising revenues. Television therefore became first available in wealthier and more densely populated areas. The fact that wealthier areas experienced the television shock first doesn't per-se pose a problem for identification. A difference in differences analysis makes use of changes within a given labor market and controls for for time invariant characteristics across labor markets. If differences in treated and untreaded areas are fixed the difference in difference estimator will be unbaised. The difference in differences analysis requires that differences between treated and untreated areas don't change for reasons other than the treatment (e.g. different trends) during the sample period. The analysis below will show that this is likely to hold in this case.

To cut through the debate about endogenous antenna location, I will present further evidence from quasi-random variation in the location of television production. This variation comes from government intervention in licensing new television stations. To launch a new antenna a construction permit had to be obtained from the FCC. The FCC had to ensure that different broadcasters signals did not interfere with each other. Construction of new antennas therefore underwent a thorough planning and licensing process. The final license put strict constraints on the time frame for the set up and the reach of the antenna. The timing and size of the market that could be reached from an antennas was thus heavily influenced by government regulation. Such regulations are likely to be exogenous to changes

[4]Exploratory broadcasting experiments, technical showcases and experimental broadcasting had taken place in a few locations and on road sows since the 1920s and had familiarized the population with the new technology. In four cities experimental broadcasters where later turned into commercial television channels. Prior to the launch of commercial television the private ownership of TV sets was however minimal.

in the local labor market for entertainers and provide quasi-random variation that will be useful in the analysis. I digitize detailed data about the application process for television antennas. This allows me to document a complete shut down in the approval of applications from 1948 until 1952. The reason for this blanked ban on construction permits were scientific constraints in transmission technology. The spectrum of frequencies on which signals can be transmitted is limited. Each broadcast blocks a certain chunk of the frequency spectrum in an area and the neighboring areas. Lack in planning in the early television days had led to a sub-optimal allocation of spectra across the country and a shortage of frequencies. The government thus decided to develop a strategic plan to re-allocate frequencies and in the mean-time issued a ban on new construction permits. This ban lasted until 1952 and left large parts of the US without TV signal at the start of the 1950s. I digitize detailed data on the application process and identify applications that were delayed due to the freeze in new permits. These stations "that didn't happen" will provide a valuable control group for the analysis.

The information about the location of television stations comes from historic industry magazines. Information on the launch date of transmission stations and their market reach was first compiled by Gentzkow (2006). This data comes from the Annual Television Factbook and is available through the ICPSR. I add to this data. For one, I add information on the location of television studio and geocode this location. Further, I digitize detailed information on the applications for new stations and geocode the location and signal reach of television antennas. This data traces which applications have been approved, which ones are pending and what antenna specifications were approved. The information comes from publications of the Federal Communications Commission as reported in the television annual supplements. This data allows me to trace the application procedure and use exogenous delays in the approval process to identify the the causal effect of a station.

Finally, I measure the market size of a given station. This is calculated based on the historic geographic area that the antenna could service. Figure 11 shows the areas reached by television signal in 1949 in blue. This information is matched with county level records of TV ownership. The combined data allows to measure the size of the audience of a given station. Since we are interested in the change in the market size, I match this information with data on performance venues. From theatrical guides of the early 20th century I digitize information on the capacity of over 3,000 US performance venues, which span more than 80% of US commuting zones. The distribution of the per-TV venue size is illustrated by Figure 12. Dark blue colors indicate larger venues. Both the venue and TV data measure the potential audience of a single show. This data therefore allows me to estimate the change in market size induced by the introduction of television.

4 Data

The data for this paper combines information from a variety of sources. A large part of the data is novel and has been compiled specifically for this analysis from historic sources. The data on television was described above. This information is combined with data from the US Census and the ICPSR.

4.1 entertainers data

The data on the labor market for entertainers comes from the US decennial census. The analysis uses the public use micro data of the US decennial census from 1920-1970. For 1920-1940 the data covers the full population of US residents and for later years a representative sample of the population. The period covers the entire roll out period and allows for substantial pre and post periods. Major changes in variable definitions prevent an extension of the time period further.[5] The data covers a representative sample of the US population in all years. [6]. Weights are used to calculate aggregate values.

I define geographic areas that constitute a local labor markets. Here I follow Autor and Dorn (2013) and group areas based on commuting patterns. The resulting areas are defined by strong within region commuting and weak across region commuting. There are 722 commuting zones covering the mainland USA.[7] On average a commuting zone has about 400,000 inhabitants and ca. 500 workers in leisure activities.

For each of the labor markets I calculate the employment in the leisure occupations and the number of top earners. Wage data is first collected in the Census in 1940 and than available consistently throughout the period. The full distribution of wages are available in 1940, but from 1950 onwards top coding applies. In all years the top code bites above the 99th percentile of the distribution. Hence incomes are fully visible except for the very top of the distribution. I therefore can't analyze top income shares or log wages of top earners. My baseline measure of top earners is the number of entertainers in the top percentile of the wage distribution. This measure is well defined in top coded data. And has a natural analogue in quantile regressions. A quantile regression analyses the wage at a given percentile. My measure is the inverse and looks at the number of people above a given wage threshold. My baseline threshold is the 99th percentile of the overall wage distribution. For each labor

[5] From 1980 onwards the Census uses different occupation groups that make comparisons with previous periods difficult. Prior to 1920 a significant change in the definition of employed workers occurs

[6] In 1940 the full count census is available, the other years are census samples

[7] These regions are consistently defined over time. Dorn provides crosswalks for 1950 and 1970. I build additional crosswalks for the remaining years.

market I count the number of entertainers in the top 1% of the overall wage distribution. In principle one could look at the number of millionaires or any other threshold. However, the number of people above an absolute wage threshold would increase mechanical if wages grow. A relative threshold from the wage distribution has the advantage that it eliminates the mechanical effect of wage growth.[8] An alternative measure is the number of top paid entertainers in a given labor market. Such a measure would for example count the number of entertainers in the top 1% of the entertainer wage distribution. This measures looks at the wage dispersion within the profession. Note that in aggregate it is mechanical that 1% of all entertainers are above the 99th percentile of the wage distribution. The variation of interest thus comes from the regional dispersion of top earners. A drawback of this measure is that the threshold might be influenced by television directly. If television causes exits or entries in entertainment the threshold would move, thus biasing our estimates. The analysis below will use a variety of top income measures. Figure 5 shows a number of different thresholds. In practice the difference is small as most of these income thresholds move together. I will show that the results are robust to different measurements of top earners.

The analyze will use two measures of employment. First, I calculate the number of entertainers per capita. I identify workers in entertainment by using occupation information. The core of the analysis will uses five occupations: Actors, athletes, dancers, musicians and entertainers not elsewhere classified. These have been selected because they appear regularly on television. IPUMS provides a consistent occupation classification based on the 1950 census throughout the sample period. It is well known that the classification of occupations changes frequently. The 1950 classification system is used throughout the sample period. In a handful of cases the Census made changes in the recording of occupations. For instance the athlete category is discontinued in 1970.[9] To take such time shifts in the occupation definition into account, all the regressions will use occupation specific time effects. For the most part entertainment occupations are well established and there is little change to their definition of the tasks performed. Unlike occupations during the era of digitalization, entertainment occupations are remain relatively unchanged over the sample period. The comparability of the occupations over time will makes it easier to interpret the estimates.

The data on workers wages, occupation and employment is available consistently for indi-

[8]In practice this is a minor problem. Year fixed effect will control for the overall number of workers above the threshold. Identification will come from the regional distribution of top earners. Moreover, top incomes were fairly constant in the 1940s and 1950s. The difference between an absolute and relative threshold is thus small.

[9]The 1950 occupation classification system was used in the 1940-1970 censi. For the years 1920 and 1930 IPUMS has mapped the original answers into the 1950 categories.

viduals over the age of 15. I hence restrict my sample to that age group. In 1920 and 1930 the data doesn't distinguish between employment and unemployment workers. I therefore work with two definitions of the workforce in entertainment. For one, I use the conventional census definition of employment, available after 1920. To work with the entire sample, I construct a second variable that counts the workforce as the sum of employed and unemployed workers in a given occupation. This is possible since the census records occupational information for the unemployed if they previously worked. The second measure is thus a noisy measure of economic activity in the profession. The broader definition is likely to dampen the estimated displacement effects as some workers who have been displaced continue to be classified as employed. In practice however this doesn't pose a major problem. Unemployment spells are short-lived, the sample thus only includes very few cases of entertainers who are unemployed. To make sure that this doesn't significantly bias my estimates I ran the analysis using either measure of employment when both are available. The results are very similar and confirm that this is not a major problem.

I group workers in leisure activities in three groups. Workers in performance entertainment, workers in interactive leisure activities and workers in drink & dining. Performance entertainers experience a shift in their production technology due to TV while the other two groups don't. Performance entertainers are therefore the main group of interest in the analysis. This group consists of five occupations that appear regularly on TV: Actors, athletes, dancers, musicians and entertainers not elsewhere classified. The last group is relevant because it includes most circus and vaudeville acts, one of the most important forms of entertainment at the time. [10].

5 Empirical analysis

The aim of this section is to test the effect of the growth in the market size on the labor market. I will focus on two outcomes of interest: The effect on top incomes caused by an expansion of markets size. Second, the effect on employment due to the added competition. The idea here is that while technology allows some actors to grow, others will be pushed out of the market. The two effects thus respectively capture the effect of the technology on the top and bottom of the distribution. Identifying the effect on top income requires variation in the available market size. Introduction of the local TV studios will generate changes in

[10] I confirm this using the original job title given by respondents. The group includes mainly acrobats, clowns, animal trainers etc. Unfortunately the original job-title is not available for all sample years and thus doesn't allow me to divide entertainers not elsewhere classified further.

the available market size. To identify the displacement effect on the other hand one needs exogenous variation in the exposure to competition from the new technology. This will be generated by changes to TV signal reach.

The effect of television is also visible in the aggregate figures on inequality. Prior to television entertainers were already a profession with a large fraction of top earners. Figure 1 shows that around 5% of actors and athletes were in the top percentile of the overall wage distribution. This fraction shoots up sharply around the time of the introduction of television. Piketty showed in a number of studies that the 1940s and 1950s were a period of relatively stable top income shares (e.g. Piketty and Atkinson (2006)). This can also be seen in figure 2. This figure shows the gap between the median and the 95th percentile of the wage distribution in the US Census. For the overall distribution the gap remains fairly stable. This contrasts sharply with the entertainment sector. For entertainers wages at the 95th percentile grew substantially faster than median wages during the 1940s. Figure 3 shows the distribution of real wages within entertainment in 1940 and 1970. The distribution has become visibly more polarized over this time period. The aggregate figures therefore suggest that the market size growth induced by TV may have had a substantial effect on the wage distribution.

5.1 Effect on top incomes

This section will estimate the causal effect of television on top earnings. The baseline estimates will make use of the difference in differences setting that emerges due to the slow roll-out of television production. The first commercial television channels were launched in 1941 but by 1950 not even halve of US commuting zones had access to TV signal. I observe each of the 722 labor markets r over the 4 decades covered by the census. This will allow me to compare top earners within each labor market over time as well as similar areas with and without TV production in each time period. I can therefore control for time and location specific fixed effects. Effects that are specific to a market or a given time period don't pose a problem for the identification. The identifying assumption of the difference in difference setting is that the timing of TV introduction is unrelated to top income trends in the local labor market. The primary driver for the location of a new TV station is access to a market with high purchasing power and high population density. These characteristics tend to be fixed over time and thus don't threaten the identification assumption. Below I will show more formally that the timing of TV introduction is unrelated to time varying factors. The difference in difference analysis therefore provides a causal estimate of the effect of television on employment.

As outlined above, production initially takes place locally in the neighborhood of the broadcasting antenna. This changes with the invention of the VCR. At this point the choice of a production locations is unconstrained. To avoid endogeneity problems that arise from this unconstrained location decision, I will estimate the causal the effect based on the time period when the production location is constrained to take place next to the broadcasting antenna. I will later exploit quasi-experimental variation in the location of television antennas. The treatment variable is thus given by the interaction of a dummy that takes the value 1 during the era of local television P^{II} and the dummy that takes the value 1 if a broadcasting antenna operates in region r at time t $TV_{r,t}$. The broadcasting antenna only has an effect on top incomes when it brings television production with it. It loses that effect after the adoption of the VCR in 1956. Wage data in the Census is available decennially since 1940. The difference in difference equation is thus given by:

$$y_{ort} = \theta_{o,t} + \beta_r + \alpha TV_{r,t} \times P_{II} + \pi X + \varepsilon_{ort}$$

where y_{ort} measure the number of top earners in occupation o, region r and time t. By running the regression at the occupation level I can control for potential time fluctuations in the definition of the occupation coding. The standard errors $\epsilon_{l,o,t}$ are clustered at the labor market level, hence running the analysis at the occupation - labor market level won't artificially lower my standard errors. $\theta_{o,t}, \beta_r$ respectively capture region and occupation-time specific effects. The vector X is a vector of control variables. All specifications control for the relocation incentives after the invention of the VCR. This is captured by a proxy for local production cost.[11] In the robustness checks I will allow for a range of other controls

The baseline specification follows much of the literature on top earnings in focusing on a relative earning metric. Top earners are measured as the share of entertainers in the top 1% of the overall wage distribution. This relative metric allows me to side step problems that may arise from growth in productivity or due to top coding. The robustness checks explore a number of alternative measures for top earners.

Table 1 reports the results. The introduction of TV production in an area leads to marked increase in the number of top earners in an area. The number of people in the top 1% increases by 90. Given an average count of 110 this implies the number of top earning entertainers almost doubles. The main threat to the identification is that the introduction of TV is related to time varying factors that also affect employment in entertainment. As a first step I add demographic controls that may capture some of the time varying factors

[11] As proxy for the production cost, I use the share of movie productions in an area in 1920, interacted with the date of the VCR invention. This measure is pre-determined and therefore unrelated to contemporaneous changes in ε.

that one might be worried about. I control for median wage, median age, share of blacks and share women in the market. Adding these controls has very little effect on the estimate. To understand better whether the timing of television coincides with the observed effect, I allow for market specific time trends. This is a very demanding specification as it adds a large number of coefficients to the specification. Accordingly the standard errors increase. The estimate remains very significant and differential regional trends clearly don't drive the results. The introduction of television has a highly significant and economically large effect on top incomes.

This result is robust to various measures of top income. The regression above uses the count of people above the 99th percentile of the wage distribution. Such a regression can be thought of as the inverse of a quantile regression. The quantile regression estimates the wage of the person at the 99th percentile. The advantage of the specification is that it can be performed with data that contains top coding. The count of people above the 99th percentile is well defined as long as the top code bites above the 99th percentile. I show that the results are robust to various other ways of measuring top incomes. Figure 5 illustrated the evolution of various top income measures. The figure shows the evolution of the 95th and 99th percentile of the overall wage distribution as reported by Piketty and Saez (2003), the 99th percentile of wages in the US Census and the 95th percentile of the entertainer wage distribution. Broadly these measure move similarly. Table 2 repeats the above analysis using various other top income measures. Column 1 repeats the baseline estimate. Column 2 uses the number of people in the top 1% as defined by Piketty and Saez (2003). The result remains very similar to before. The number of entertainers in the top percentile doubles due to television.

Column 3 and 4 focus on the within entertainer wage distribution. These specifications analyze the regional dispersion of entertainers above the 95th percentile of the entertainer wage distribution. By definition 5% of entertainers will earn wages above the 95th percentile in aggregate. With this definition of top income, TV can not change the number of top earners. Instead it can change where these individuals live. If TV had a positive effect on top incomes, one would expect that the number of top earning entertainers increases in areas where TV productions are filmed. The analysis focuses on entertainers above the 95th percentile. With the Census data it is not possible to analyze higher percentiles, such as the 99th percentile of the entertainer wage distribution. Since the entertainer wage distribution is very skewed, the 99th percentile lies above the Census data top code. Analyzing within entertainer wage dispersion has the appealing advantage that it is a measure of inequality in the affected sector. This measure is however problematic if TV induces substantial exit in the entertainment sector. Such selection would make the 95th percentile endogenous to the introduction of television. If television results in an exit of the bottom 10% of entertainers, the 95th wage percentile would rise. If there was no further effect on top earners, we would find that fewer entertainers are top earners after the introduction of television. Hence, this measure will lead to a downward biased in the estimate of TV. Indeed in column 3 the number of top earners increases by less. The increase here is 30%. To address the endogeneity issue column 4 keeps the threshold fixed at the 1940 level. This measure is thus unaffected by exit of entertainers. The estimate is indeed substantially bigger than column 3. These results confirm that television led to a substantial increase in top earnings in entertainment.

The identifying assumption of the difference in difference regression is that differences between treated and untreated groups stay constant in the absence of the treatment. This assumption is not directly testable. A useful safeguard is to ensure that there are no pre-trends. The idea of this test is to rule out that different trends in treatment and control group lead to a spurious estimate. If such trends continue post treatment one would find differences between treatment and control group even if there is no true treatment effect. The timing of the effect of treatment thus helps us check that we are indeed identifying the causal effect of the treatment. The introduction and removal of local TV production provides a strong check for such trends. We can check both the pre treatment period and the post treatment removal periods. If there are common trends we would expect that after the removal of the treatment the differences between treatment and control group revert to the pre treatment differences.[12] This therefore allows to go beyond the standard pre-trend check and provide an even more powerful test for common trends throughout the treatment period. We don't rely only on the pre-period to extrapolate trends but can check differences in treatment and control group before the treatment and after the treatment removal. This is possible since we observe three periods: the period before, during and after the removal of the treatment. We would therefore like to see that the treatment effect occurs when the larger market size was available and disappears thereafter. Figure 6 illustrates the timing of the effect of bigger market size. The number of top earners rises when television production is introduced. Recall that the invention of the videocassette recorder should make the location of broadcasting antennas irrelevant for production Indeed the figure shows that the effect of having a TV antenna fades when local television production starts to disappear. By 1969 the differences between treatment and control group returned to the pre treatment level. Differences between treatment and control group thus are the same before and after treatment removal. There is no evidence of differential trends. This suggests that differences

[12]This requires that the temporary increase in available market size has no lasting effect.

between treatment and control group during the era of local TV production is indeed caused by television.

For peace of mind I provide the conventional pre-trend checks as well. As argued above relying purely on the pre-treatment period is a weaker check. Nevertheless, I report the pre trends in table 4. An added challenge for estimating pre-trends is that wage data in the Census is available from 1939 onwards. This only gives me a single pre-treatment period. To estimate pre-trends I therefore combine the Census data with data from Internal Revenue Services (IRS) tax return data. In 1916 the IRS published aggregate information on top earners by occupation-state bins. Data for actors and athletes are reported. I link the Census data with the tax data and run the regressions at the state level. Column 1 repeats the baseline estimate with data aggregated at the state level. The effect is also highly significant at this aggregated level. Column 2 adds the additional 1916 data from the IRS. The results stay unchanged. Column 3 shows the differences in top earners in treatment and control group for the various years. It shows a marked jump up in top earners in the treated group in the year of local TV production. The coefficient on the pre-trend is not significant because the standard errors are large. If anything the pre-period points to a decrease in top earners in the treatment area. Even if taken at face value the pre-trends thus can't explain the identified effect.

The results this far make a strong case that the television effects can be interpreted as causal. The robustness checks rule out that the results are driven by fixed and a number of variable characteristics of labor markets. One point that the analysis hasn't address so far is that the introduction of television could coincide with other location specific changes in the labor market. To rule this out I will perform a number of placebo tests. This will confirm that the effect of television is affecting only the types of tasks we would expect it to affect. An ideal placebo test analyses the change in top incomes for a group of workers who weren't affected by television for exogenous reasons but who are otherwise similar. Television only allows certain leisure occupations to expand their market. Many leisure activities are by nature interactive. Activities such as bowling, sports lessons or dining out can't be performed through television. Two groups of leisure workers who don't experience a market size change from television are interactive leisure occupations and drink & dine occupations. These groups do however share effects common to the entertainment sector. I run a placebo test on these groups. Figure 7 reports the effect of television. Indeed no growth in top incomes occurs for these workers. A second placebo group are other top paying occupations. This allows to check whether the effects are driven by location-specific changes in returns to skill. Table 3 reports the effect of television for various other high paying occupations. Only performance entertainers experience the significant and large top earner rise when television production occurs in a given area.

Column 3 runs a full triple difference regression. Using the placebo occupations I have treated and untreated workers within each labor market. We already controlled for location specific trends before, this specification will go further and allow for a non-parametric location specific time fixed effect. An example where this might be necessary is if improved local credit conditions result in greater demand for premium entertainment and simultaneously lead to the launch of a new TV channel. This may lead to an upward bias in the estimates. My treatment now varies at the time, labor market and occupation level. This allows me to control for pairwise interactions of time, market and occupation fixed effects. These will address the outlined credit access problem as the fixed effects will now absorb location specific time effects. The specification thus looks as follows:

$$y_{ort} = \delta_{o,t} + \theta_{l,o} + \phi_{l,t} + \alpha TV_{r,t,o} \times P_{II} + \pi X + \varepsilon_{ort}$$

with δ, θ, ϕ representing fixed effects, also note the change in subscripts on the TV variable. Again I find that the effect of TV studios is concentrated on the entertainment performers. It is reassuring that the point estimates remain similar to the baseline estimates. The introduction of a "superstar technology" thus has a large causal effect on top incomes and this effect is unique to the treated group.

The difference in differences analysis doesn't require random assignment of the treatment. Nevertheless, quasi-random variation will allow us to check that the introduction of television is not spuriously related to changes in top incomes. In this setting a government interventions provides quasi-random variation in the treatment. The government freezes the issue of new construction permits from 1948 to 1952 due to the signal problems described above. It gives rise to a number of regions that would have received television at a similar point in time, but due to the freeze get access years apart.

The timing of the the introduction of television is determined by a government approval process. An entrepreneur who wants to set up a TV station needs to file an application with the FCC. The FCC then analyses the effect the new broadcasting aerial would have on existing infrastructure. In particular it considers what frequency spectra are available and what interference is likely to occur. TV signal in the early day was broadcast via airwaves. To be able to broadcast the no other signal can be send on that frequency. The FCC had to safeguard that there is a unique user for a frequency. Frequency spectra are thus a rival good. Television signal could only be broadcast on a narrow range of frequencies. There thus was a natural limit to the number of sending stations. To complicate things further frequency spectra were also used by a variety of non TV actors, such as civil aviation, emergency services, military etc. Moreover, if a spectrum was assigned to an actor in one area, none would be able to use this frequency in surrounding areas either to avoid interference at the fringes. The FCC thus faced a complex assignment problem. I digitize data from FCC records on the application process for aerials. Throughout the 40s demand for TV stations was growing rapidly. Towards the end of the 1940s it became clear that the FCC assignment system would not be able to keep up with demand. The FCC commissioned a rethink of the assignment system and banned the issue of new construction permits in the mean time. In 1948 a large number of applications were delayed. Figure 9 shows the number of approved applications over time. The ban resulted in a sharp drop in the number of new permits. This ban on new stations lasted until 1952. The government intervention therefore delayed the introduction of television stations by several years in a number of labor markets.

I use the quasi-random variation introduced by the ban in the next regressions. One might be worried that labor markets for performance entertainers are somehow different in areas that apply for TV permits. The exogenous ban on construction permits allows to address this worry. I will run a placebo test with the station that "didn't happen" to confirm that the effect only appears if TV productions are taking place. The fact that an area is applying for a TV station doesn't reveal something that is spuriously related to top income changes. Figure 9 shows the approved stations in green and the stations that fell under the ban in red. I run the baseline regression with the stations that "didn't happen" as treatment group. Figure 10 shows the effect of stations that "didn't happen." The effect of such stations is a precise zero. Labor markets that applied for licenses but didn't get a station for exogenous reasons thus experience no change in top earners whatsoever. This confirms that top wages only grow if TV production is actually launched in an area.

To answer the question how much top earners gain from a given percentage change in market size it will be necessary to quantify the effect of television on market size. I therefore run the baseline specification with market size as outcome variable. The market size variable measures potential audience of a single performance in a given labor market. It makes use of detailed information on the capacity of over 3,000 performance venues throughout the US. This data has been digitized from theatrical guides from the first halve of the 20th century. This data is combined with information on television ownership within the reach of each of the local television stations. I thus can measure how many people an entertainer could reach with a single performance on stage or TV. Table 5 reports the results of a regression of the log of audience size on television. The arrival of television leads to 1.4 log points growth in audience size. Converting the log points this implies a market size growth by factor 4 to 5.

Combining the estimates we can calculate the elasticity of top earner to market size. This is an "IV estimate" using the timing of TV arrival as instrument. The effect of television on top earners is the reduced form and the effect of television on market size the first stage. Dividing the two estimates will give an IV estimate for the effect of market size on top earners. The implied elasticity is around 0.6. A 10% growth in the available market size thus results in a 6% increase in the number of top earners.

This parameter is closely linked to the parameter of the Gabaix & Landier model discussed above. The relevant structural parameter is the elasticity of top wages to market size. The estimate discussed so far instead is the elasticity of the number of top earners. The CDF function of wages maps a change in wages to a change in the number of top earners:

$$TE = (1 - F(w^{99}))$$

$$\triangle_{TE} = f(w^{99})\triangle_w$$

If the shape of the CDF is known this equation allows to translate a change in employment to an associated shift in wages. The shape of the tail of the wage distribution has received much attention in the literature. In Kuznet's pioneering study in 1953 he finds that the US wage distribution can be well approximated by a Pareto distribution. An overwhelming number of studies have since replicated this finding for a variety of time periods and countries. A convenient property of the Pareto distribution is that the tail of the distribution has a well defined shape with $\frac{f(x)}{1-F(x)} = \alpha/x$. Using this fact we can re-write the above equation in terms of elasticities $\varepsilon_{i,j}$ with α the pareto coefficient:

$$\frac{\triangle_{TE}}{TE} = \frac{f(w^{99})\triangle_w}{(1 - F(w^{99}))}$$
$$\varepsilon_{TE,m} = \alpha\varepsilon_{w,m}$$

This gives us a simple expression for the link between the elasticity of quantity and prices to market size. I estimate the α parameter on the pre-TV distribution using the full Census of 1940. I experiment with a number of estimation strategies with similar results.[13] Independent of the approach the coefficient is close to but bigger than 3. To err on the conservative side, I will use a value of 3 for the analysis.

In the analzsis above we found $\varepsilon_{w,m} = 0.57$. Using the relation derived here we can translate this into an elasticity of income. The implied eslasticity of top wages to market

[13] The baseline results use Kuznet's approach to estimate the pareto parameter. This approach uses the fact that average income above a threshold is proportional to the threshold. With α the coefficient of proportionality. I also run the Atkinson & Piketty approch and use different threshold values. All estimates are above 3, with most between 3.02 and 3.16.

size is $\varepsilon_{w,m} \approx 0.19$. A doubling in market size will thus raise top wages around 20%.

The elasticity estimate of 0.19 is substantially smaller than the correlation of market size and top wages. OLS estimates of top wages on market size find a coefficient around 0.33 ($\hat{\beta}^{OLS} = \frac{1}{3}$). The IV estimate thus implies that the OLS estimates overstate the superstar effect by about 40%.

It is perhaps unsurprising that the OLS estimate is upward bias. Such regressions don't controll for the potential endogenous change in market size. It seems likely that omitted factors like effort are positively correlated with market size growth and wages. Using exogenous variation in market size the IV estimate instead reflects the causal effect of market size on wages at the top. This causal effect is smaller than previously thought and consequenlty "superstar effects" explain less of the rapid growth in top incomes than many economist hypothesist.

5.2 Effect on employment

To complete the analysis of the "superstar technology" on inequality I will now look at the effects at the bottom of the distribution. Television increased the competition for many entertainers in local theaters and circuses. To track the introduction of competition I will use slightly different variation to the previous section. Competition from television occurs if TV signal can be received in that area. Figure 11 shows the extend of signal reach of TV stations in 1950 in blue. Depending on the terrain and the technical features of the antenna the signal reach can vary from a few miles to around 100 miles. Signal is difficult to target and many areas that weren't aimed for became "treated" nevertheless. The set of areas that experience greater competition from technology are thus very diverse. Using the variation in signal reach I estimate the displacement effect of television. Column 1 of Table 6 shows that the introduction of TV signal has a significant negative effect on employment of around 2 entertainers per 100,000 inhabitants. The average of employment is around 9, hence TV signal results roughly in a 20% drop in employment. Columns 2 and 3 include demographics and location specific trends and confirm that the results are robust.

Next I make use of the stations that "didn't happen." I calculate which areas would have experienced competition had the government not issued a ban of new television antennas. These areas are colored in red in Figure 11. This data is only available for the period prior to 1960 and the sample is reduced accordingly. Column 4 shows that there is no employment loss in areas that receive signal that "didn't happen". Next, I leverage the fact that signal travels beyond the area were the antenna is set up. Some areas receive signal despite not applying for TV antenna licenses. This occurs because TV signal travels beyond the area of origin. Arguably, the introduction of treatment is more exogenous if an area didn't apply for a TV antenna. Column 5 shows that the effect on these areas. The effects are in the same ball-park as the baseline estimates. Finally, column 6 narrows the sample further and only compares areas that received signal with areas that would have received TV in the absence of the freeze. These are the blue and red areas in figure 11. The estimate shows that employment falls in areas that received signal relative to areas that didn't because of the government ban on new antennas.

Finally, I compare the effect of television across different entertainment occupations. Television only provides a close substitute for some leisure activities. I use the same two control groups as above: drinking and dining and interactive leisure activities. Of course we would expect these occupations to also suffer from the introduction of television, but they should be less substitutable by television. Hence, we would expect smaller negative effects on these occupations. I run the same difference in difference analysis for workers in these occupations. Table 7 reports the estimates. The employment loss is largest among performance entertainers but also occurs among the other professions. This confirms that the effect of television is indeed task specific and particularly strong for occupations that are close substitutes to television. The last column performs a full "triple difference analysis." This analysis adds the untreated leisure activities from above to the specification and controls for pairwise interactions of fixed effects. The estimate shows significantly negative employment effects of -6. If anything the point estimates is bigger than in the baseline specification. This may suggest that the baseline underestimates the true loss in employment. However, the standard errors are large here. The baseline estimate is well within the confidence interval of the triple difference estimate. The causal effect of television on employment is therefore in the order of 20% of employment.

6 Conclusion

There have been vast changes in top incomes over the last decades but little is known about the causes for this. Companies have gone from small local enterprises to global operations. Technological innovation, foremost in communication technologies have made it easier to operate over vast distances. These changes have been associated with growth in top incomes. This paper is the first to provide causal evidence on the effect of growing scalability of production.

The paper analyses such "superstar effects" in the entertainment industry. The staggered introduction of television has led to a substantial growth in the market reach of entertainers. A single performance has been able to increase his audience by factor 4 with the help of the new technology. Regulatory and technological problems led to quasi random variation in the roll-out of the technology and allows me to identify the causal effect. I find that a doubling in the market size leads to at 60% increase in the number of top earners in entertainment. Translated into an effect on top wages, the estimate implies an elasticity of wages at the top of around 0.2. A value well below the correlation of market size and top wages found in many data sets. The IV estimate thus suggests that an OLS regression of top wages on market size is upward bias. The literature has therefore likely overplayed the importance of the "superstar effect" for top income growth. The superstar technology has also created losers. The ability to offer one's product simultaneously to a big audience pushed a number of workers out of the market. I find that competition from the new technology leads to a reduction in employment of ca. 20%. The increase in production scalability has therefore profound effects on inequality both at the top and bottom of the distribution.

7 Appendix

8 Data Details

8.1 Geography

- I use 1990 CZ as geographic units of the mainland US
- Dorn crosswalks are used for the 1950 and 1970 data
- for 1960 no crosswalk is provided. I match the 1960 PUMAs to 2000 pumas and aggregate at the CZ level that way. Luckily there is almost no change in the PUMAs between 1960 and 2000
- to crosswalk the 1940 data, I build a crosswalk based on the overlap of 1940 counties and 1990 CZ. In most cases counties fall into a single CZ. A handful of counties are split between CZ. For cases where more than 3 percent of the area falls into another CZ, I construct a weight that assigns an observation to both commuting zones. The two observations are given weights so that they together count as a single observation. The weight is the share of the county's area falling into the CZ. The same procedure is followed for 1960 mini PUMAs
- Carson city county (ICSPR 650510) poses a problem. This county only emerges as a merger of Ormsby and Carson City in 1969, but observations in IPUMS are already assigned to this county in 1940. I assign them to Ormsby county (650250)
- czone 28602 has no employed individual in 1940 and no wage record.

8.2 Worker data

The data uses the public use micro data of the US decennial census from 1940-1970 (excluding Hawaii and Alaska). The data is aggregated at the commuting zone level as defined by DORN (Autor and Dorn (2013))

- there are 722 commuting zones (CZ) covering the mainland USA. These regions are consistently defined over time.
- there are 44 relevant occupations
- controls are population aggregates in the area: share high skilled (high school and above for people over 25), share non white, median age, sample size per CZ, median wage and age
- Aggregates are calculated using the provided sample weights
- variables used incwage, occ1950 (in combination with empstat), wkswork2, hrswork2

8.3 Employment

- Occupation based on the 1950 classification of IPUMS (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. Shifts in the definition over time will be captured by the occupation specific year FE in the analysis. Combining data from repeated sample years requires consistency across years. IPUMS provides a detailed methodology to achieve close matches across various vintages of the US census. Luckily the occupations used in this analysis are little affected by changes over time. More details on the changes and how they have been dealt with are: The pre 1950 samples use an occupation system that IPUMS judges to be almost equivalent. For those samples IPUMS states: "the 1940 was very similar to 1950, incorporating these two years into OCC1950 required very little judgment on our part. With the exception of a small number of cases in the 1910 data, the pre1940 samples already contained OCC1950, as described above." For the majority of years no adjustment all is therefore necessary. Changes for the 1950-1960 period - Actors (1950 employment count in terms of 1950 code: 14,921 and in terms of 1960 code: 14,721), other entertainment professions are unaffected. Changes from 1960-1970: Pre 1970 teachers in music and dancing were paired with musicians and dancers. In 1970 teachers become a separate category. My analysis excludes teachers and thus is unaffected by this change. Athletes disappear in 1970 coding. The analysis therefore only uses the athlete occupation until 1960. The only change that has a major effect on worker counts is for "Entertainers nec". In 1970 ca. 9,000 workers that were previously categorized as "professional technical and kindered workers" are added and a few workers from other categories. The added workers account for ca. 40 percent of the new occupation group. The occupation specific year effect ought to absorb this change. I have also performed the analysis excluding 1970 and find similar results. Moreover I find the TV effects for each occupation individually. The classification changes therefore seem to have little effect on the results.
- The industry classification also changes over time. I use the industry variable to rule eliminate teachers from the occupation categories "Musicians and music teacher" and "Dancers and dance teachers." The census documentation does not note any significant changes to the definition of education services over the sample period, however the scope of the variable fluctuates substantially over time. From 1930 to 1940 the employment falls from around 70,000 to 20,000, from 1950 to 1960 it increases to around 200,000 and falls back to around 90,000 from 1960 to 1970.

- I use two types of control groups. For one, entertainers whose production technology is unchanged by TV (these are drink & dining professions and interactive leisure occupations). Second, top earning professions outside entertainment (lawyer, medics, engineers, managers, financial service). The relevant occupations are available across most years. Exceptions are 1940 where a few occupations in engineering, medicine and interactive leisure are grouped together and in 1970 where the floor men category is discontinued. I control for those changes with year-occupation fixed effects in the regressions. To demonstrate that these changes don't drive the results, I show all results for groups affected and unaffected by changes separately. Neither of them react significantly to the introduction of TV.
- Number of workers are based on labforce and empstat. Both variables are consistently available for 16+ year olds. Hence the sample is restricted to that age group.
- weeks worked is based on mid-points of intervalled weeks worked. This variable is available consistently over time
- occupation is recorded for age>14. I use this information for all employed. This is available consistently with the exception of institutional inmates who are excluded until 1960. The magnitude of this change is small and the time fixed effect will absorb the effect on the overall level of employment.
- weeks worker in 1940 refer to "full-time equivalent" weeks, refers to previous calender year, in 1950 only sample line individuals are asked about weeks
- hours data is comparable over years, refers to previous week

8.4 Earnings data

- labor earnings are used to be consistent with the model (wages, salaries, commissions, cash bonuses, tips, and other money income received from an employer) - should do consistency check with total income at some point...
- census variable refers to the previous calender year
- In 1940 and 1960+ every individual replies to this question - in 1950 only sample line individuals
- Wage data is in real 1950 terms
- Top earners are individuals above the 99th percentile of annual earnings who report positive earnings. The 99th percentile threshold is always below the top code, hence the top code doesn't pose a problem here.
- The 1940 100% sample is not top coded. For the smaller sample the top code is a bit iffy in 1940 when the 99th percentile is only 1 Dollar below the top code)

This differs from Piketty et al who use earnings data of tax units. As described above, I use wage data and focus on individual data rather than earnings of a tax unit. This choice makes economically sense for this setting. The superstar theory is concerned with individual labor earnings and abstracts from household composition and capital income.

- As a robustness check I use earners above the 99th percentile within their occupation.

8.5 Controls

- Control variables are: share blacks, male, high skilled and median age and income. Most variables are available consistently throughout the sample period. Income and education are only available from 1940 onwards. The race variable as has changing categories and varying treatment of mixed race individuals. I use the IPUMS harmonized race variable that corrects for those fluctuations were possible.

9 Figures

Figure 1: Share Actors & Athletes in Top 1% of Wage Distribution

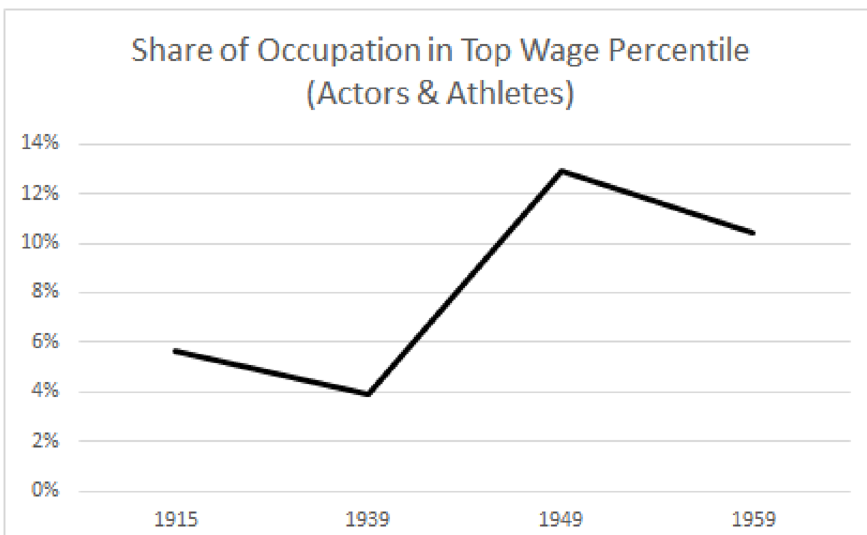


Figure 2: P95-P50 Gap

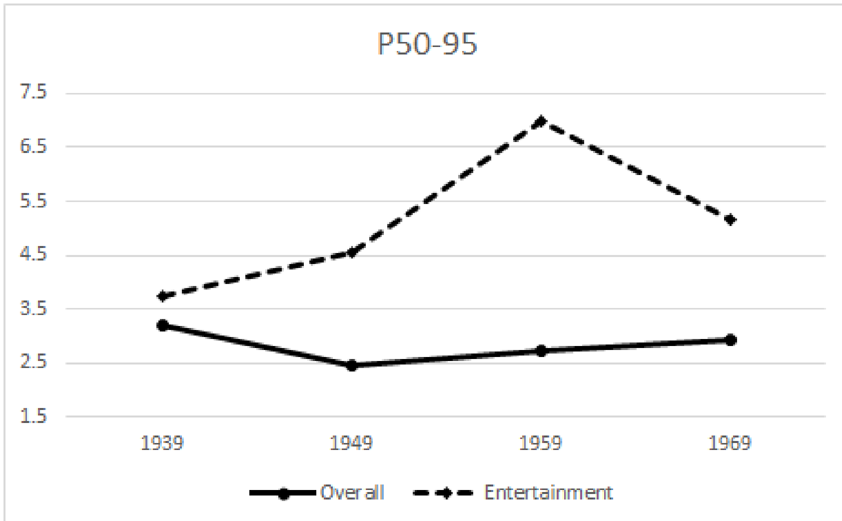


Figure 3: Wage Distribution 1940 and 1970

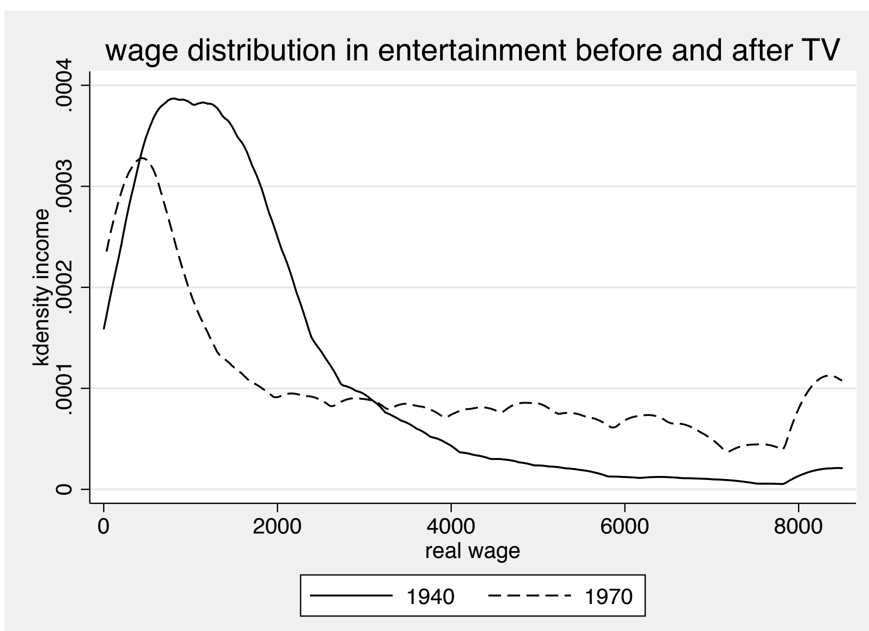


Figure 4: TV Production 1950

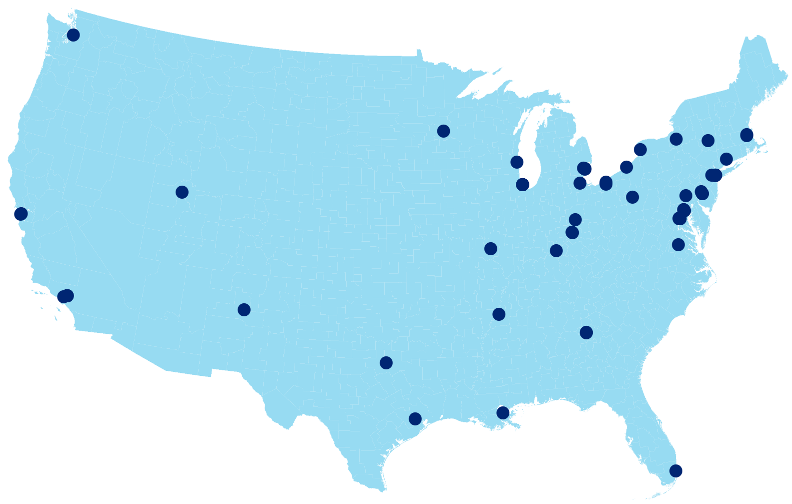


Figure 5: Top Income Percentile Values

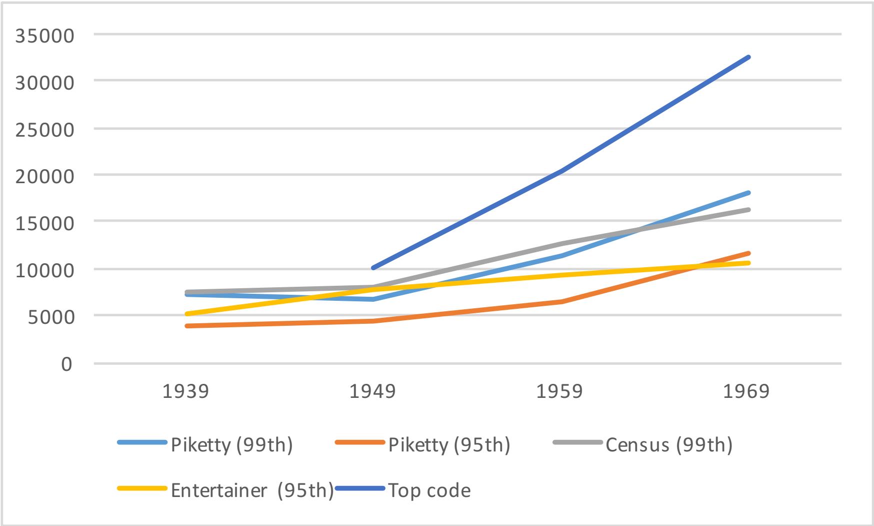


Figure 6: Effect on Performance Leisure Occupations

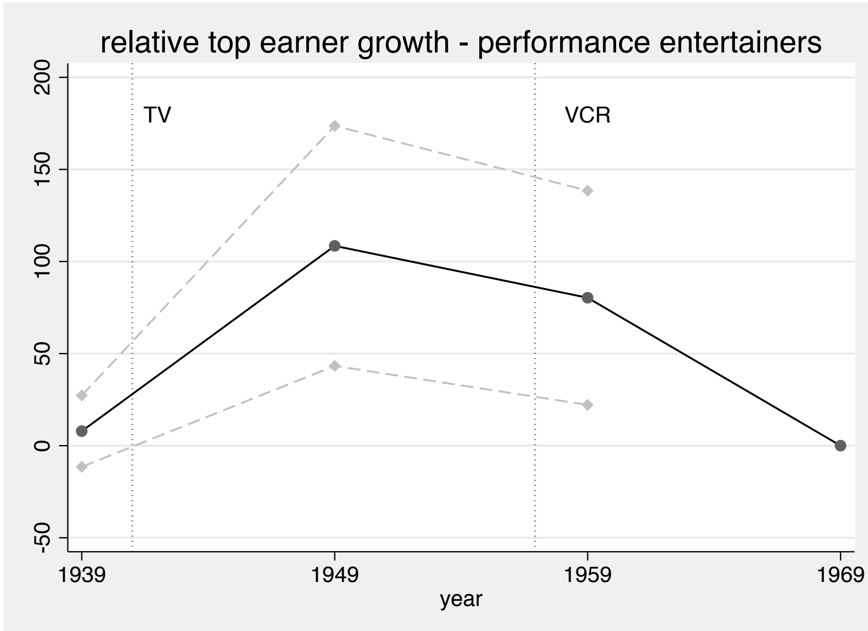


Figure 7: Effect on Interactive Leisure Occupations

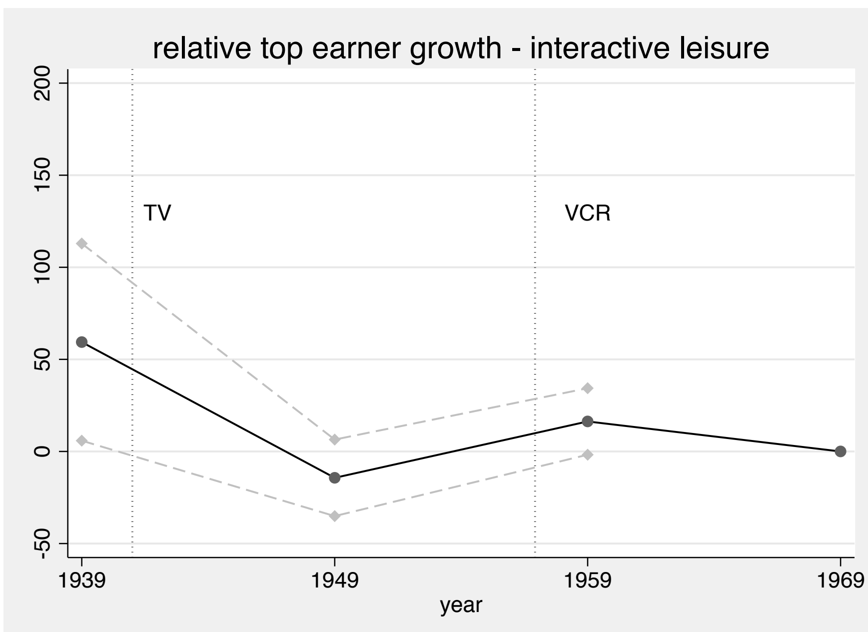


Figure 8: Number of TV Licenses Granted

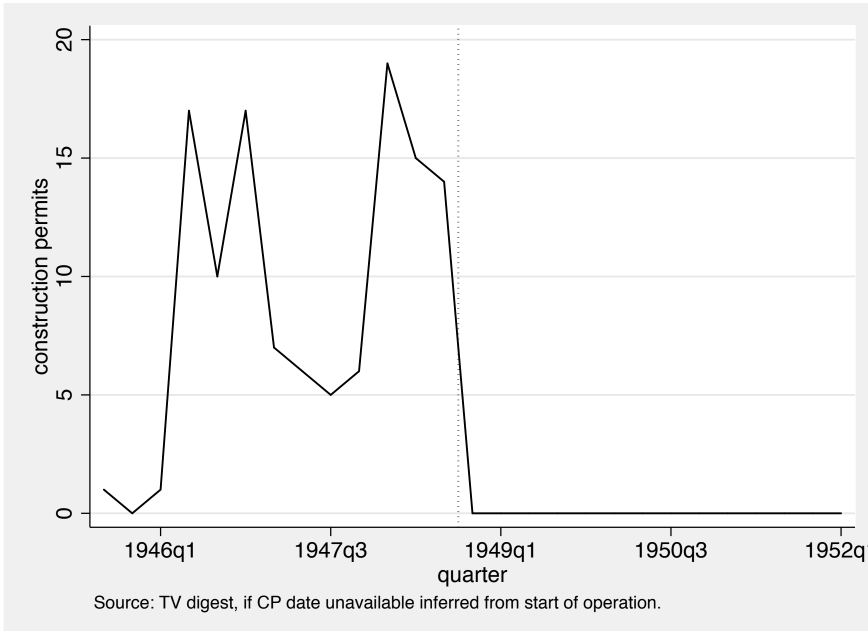


Figure 9: Granted and Frozen TV Licenses

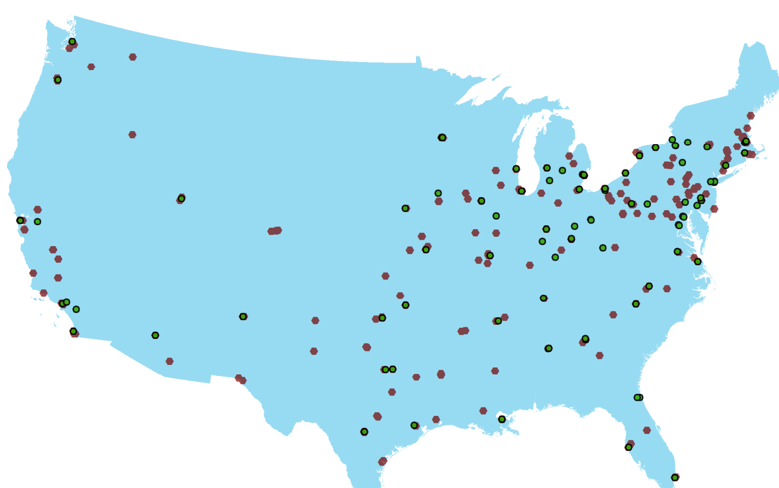


Figure 10: Effect on Placebo Stations

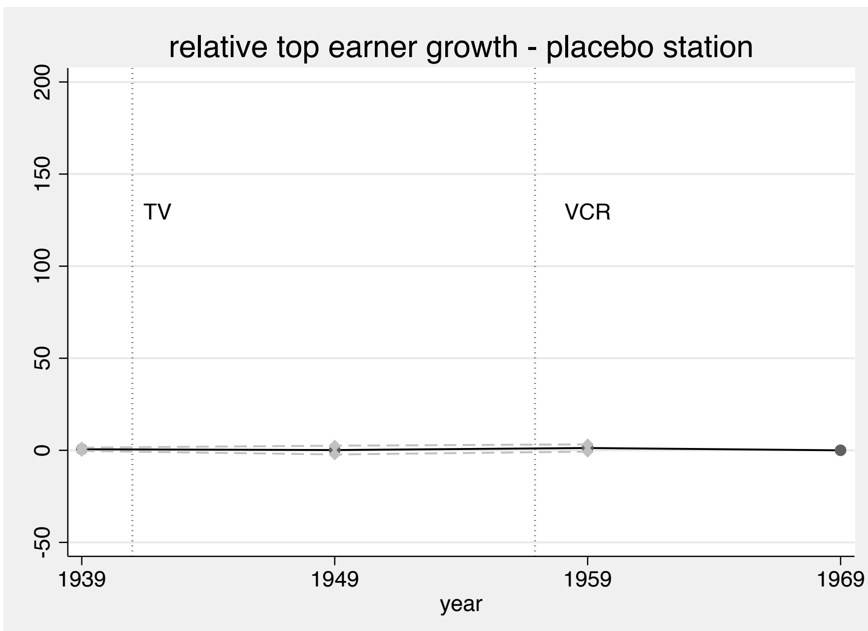


Figure 11: Signal Reach of Licensed and Frozen Stations

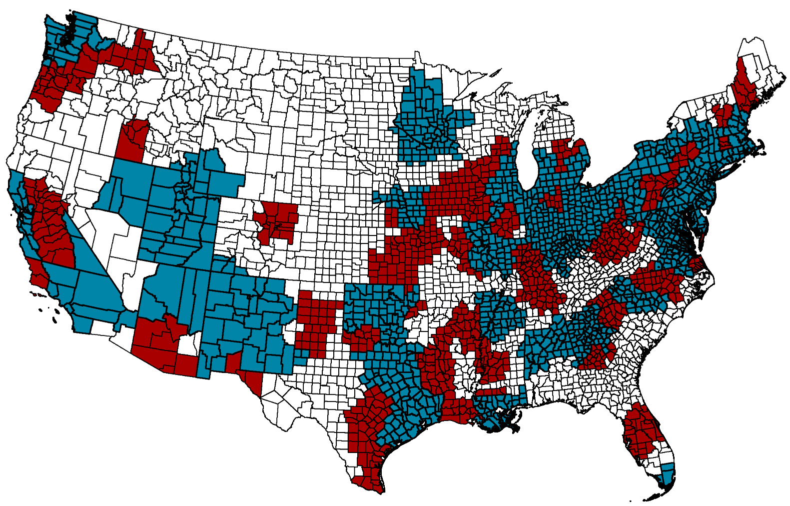
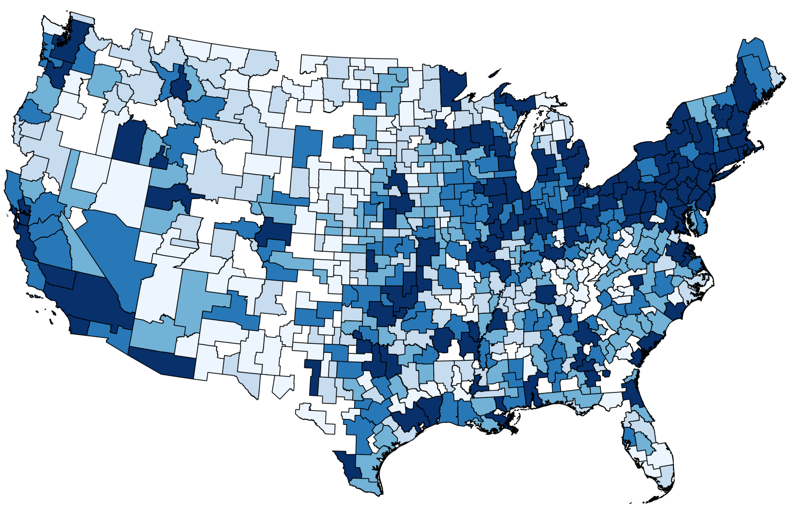


Figure 12: Size of Performance Venues



10 Tables

Table 1: Effect on top earning - instrument

	(1)	(2)	(3)
TV X P_{II}	90.19	93.16	127.6
	(26.25)	(26.16)	(47.68)
demographics		✓	
commuting zone trends			✓
mean LHS	110.62		

Source: Data US Census 1940-1970. Dependent variable is the number of workers in the top 1% normalised by employment in the occupation. The number of observations are 13718. All regressions controll for commuting zone fixed effects, year-occupation fixed effects and film amenities after 1956. Standard errors are clustered at the commuting zone level. Demographics are median age, income and share female, minority and population density and different trends for urban areas.

Table 2: Within Entertainment dispersion

	(1)	(2)	(3)	(4)
TV X P_{II}	90.19	132.5	31.64	120.0
	(26.25)	(35.92)	(16.36)	(47.85)
mean LHS	110.62	158.14	110.62	205.72

Source: Data US Census 1940-1970. Dependent variable is the number of workers in the top 1% normalised by employment in the occupation. The number of observations are 13718. All regressions controll for commuting zone fixed effects, year-occupation fixed effects and film amenities after 1956. Standard errors are clustered at the commuting zone level. Demographics are median age, income and share female, minority and population density and different trends for urban areas.

Table 3: Earning Effect - triple diff

	(1)	(2)	(3)
treated=0 X			
TV X P_{II}	–13.37		
	(6.774)		
treated=1 X			
TV X P_{II}	116.4	116.4	94.41
	(44.94)	(44.94)	(29.32)
interactive leisure profession X			
TV X P_{II}		–60.19	
		(21.02)	
drink & dine profession X			
TV X P_{II}		–71.34	
		(33.24)	
finance, accounting, law X			
TV X P_{II}		74.61	
		(73.33)	
medics X			
TV X P_{II}		–61.53	
		(27.81)	
engineers X			
TV X P_{II}		–21.77	
		(22.76)	
managers X			
TV X P_{II}		50.33	
		(36.98)	
Observations	100308	100308	100308
location, occupation-year FE	✓	✓	
pairwise interaction of location, year, occupation FE			✓

Data and specification are as in 1. The number of observations are 100308. The control group are workers in non affected free time professions: drink & dining and active leisure and typical high pay professions: management, medicine, engineering, professional services

Table 4: Top earner analysis - state level

	(1)	(2)	(3)
1915 x TV			0
			(.)
1940 x TV			–9.616
			(5.948)
TV X P_{II}	20.94	20.18	–2.985
	(8.093)	(7.355)	(1.789)
1960 x TV			–9.951
			(6.167)
1970 x TV			–13.33
			(8.068)
Observations	912	1008	1008

Source: Data US Census and IRS. Data from 1916-1970, except first column which uses sample 1940-1970. In column 4 the omitted year is 1916. Standard errors are clustered at the state level.

Table 5: Audience growth

	(1)	(2)	(3)
TV X P_{II}	1.420	1.445	1.489
	(0.208)	(0.190)	(0.227)
demographics		✓	
commuting zone trends			✓

Source: Data US Census 1940-1970. Dependent variable is the capacity of the largest venue in the comuting zone. The number of observations are 2656. All regressions controll for commuting zone fixed effects, year fixed effects and film amenities after 1956. Standard errors are clustered at the commuting zone level. Demographics are median age, income and share female, minority and population density and different trends for urban areas.

Table 6: Employment effect of TV

	(1)	(2)	(3)	(4)	(5)	(6)
TV signal	–2.398	–2.243	–2.596	–2.041	–1.773	0
	(1.240)	(1.254)	(1.688)	(1.796)	(1.664)	(.)
placebo TV signal				–0.298	–0.126	1.645
				(0.935)	(0.840)	(1.697)
Observations	20938	20938	20938	14440	12000	4340
demographics		✓				
commuting zone trends			✓			
sample	full	full	full	<1960	<1960 & non-applicant	<1960, non-applicant, placebo only
number of cluster	722	722	722	722	600	217

Source: Data US Census 1920-1970. The dependent variable is entertainers per 100,000 inhabitants. Demographics are median age, % female, % black, population density, trends for urban areas and controls for TV production location

Table 7: Employment Effect - triple diff

	(1)	(2)
TV signal=1 X treated=1	–0.970	–6.271
	(2.604)	(3.673)
Observations	58482	58482
location, occupation-year FE	✓	
pairwise interaction of location, year, occupation FE		✓
s.e. clustered by	commuting zone	commuting zone

Source: Data US Census 1920-1970. Dep. var is count of entertainer in top 1% over workforce. The control group are workers in non affected free time professions: drink & dining and active leisure

References

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

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

Garicano, L. (2000). Hierarchies and the organization of knowledge in production. *Journal of political Economy 108*(5), 874–904.

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

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

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

Murphy, K. J. (1999). Executive compensation. In *Handbook of Labor Economics, Vol. 3*, Number June, pp. 2485–2563.

Piketty, T. and A. B. Atkinson (2006). *Top Incomes over the 20th Century.*

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

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

Terviö, M. (2008). The Difference That CEOs Make : An Assignment Model Approach. *American Economic Review 98*(3), 642–668.