Economic Growth and Inequality in the Mass Market Economy

David Mayer-Foulkes
CIDE

Kurt Hafner
Fakultät für International Business, Hochschule Heilbronn

Preliminary
8/6/2015

Abstract

The mass market economy is characterized by the combination of technological innovation and absorption and generates inefficient and unequal growth. Innovation-for-profits originates monopolistic competition in some sectors, while absorption occurs in competitive sectors for which innovation is inviable. Innovative sectors produce profits that concentrate on a small number of owners. Yet the economy-wide wage level is defined by a lower, average technological level. It follows that there are policies that can promote equality, efficiency and growth. An empirical study for US states from 1997 to 2011 confirms that the proportional size the large scale sector drives aggregate employment, payroll and wages, top 1 and top 10% income participation, and poverty.

1. Introduction

How can wages remain low while technological levels rise? Wages are predicted to be proportional to the technological level, both in general equilibrium models, and in models of endogenous technological change. How can income concentration reach such high levels in the context of market competition? That 85 people can own approximately as much wealth as 2.37 billion people, the poorer half of the world population, reflects on the dynamics of the market economy more than on specific persons.[1] Consider that the institutions that support markets allow for the coexistence of competition both with and without market power. Moreover, two core productive processes generate market power, innovation and large scale production. Yet at the same time millions of small firms produce in a context approximating perfect competition. To understand the industrial or mass market economy, it is necessary

[1]This is a back-of-the-envelope calculation by Oxfam (2014), based on Credit Suisse (2014) and Forbes.

to consider how production both with and without market power combine. The purpose of this paper is to explain how the interaction of mass production with technological innovation under monopolistic competition, and small scale production with technological absorption under perfect competition, explain the prevalence of both high income concentration and relatively low wages.

We model a *mass market economy* consisting of two sectors, one innovative, with monopolistic competition, and the other absorptive[2], with perfect competition. Our model constructs a new macroeconomic perspective on economic growth and distribution that results from combining the dynamics of market power and competition. We conduct a cointegration analysis of US states from 1997 to 2011. The econometric analysis tests the relevance of this perspective, and shows that key indicators of large scale production, the proportion of employment and payroll in firms with 500 or more employees, drive aggregate payroll, employment, wages, productivity, top one and top ten percent income shares, and poverty rates, as predicted by the model. To the best of our knowledge there is no previous study of the impact of the large scale sector on the economic growth of the aggregate economy, except of course in so far as R&D takes place in this sector.

The idea prevails that competitive general equilibrium is an approximate representation of the market economy. Yet at the same time, the theory of endogenous technological change is based on the idea that market power provides the incentives for innovation, and two decades of research have confirmed that economic growth is driven by technological change. These two points of view therefore need to be combined. Moreover, the role of the large scale sector is not small in the US. In 2012, 51.6% of the workforce was employed in the 0.3% of firms with 500 or more employees. Meanwhile 89.6% of the 5,726,160 firms had less than 20 employees and employed 17.6% of the workforce.[3]

It is clear that technological change consists of both innovation and adoption. The mass market economy model applies to any market economy that has an innovating, large scale production sector with market power, and a technologically absorbing, small scale, competitive sector. Thus it applies to industrial market economies in general, both developed and underdeveloped, since the Industrial Revolution developed large scale production.

While large scale production and innovation generate productivity and growth, through their market power they also generate income concentration and inefficiency.[4] The detrimental impacts of market power on efficiency and distribution are clear at the static level. What is crucial is to investigate the dynamic context, were market power also presents detrimental impacts on innovation, for two reasons. First, the input mix of innovative and backward

[2]We refer interchangeably to technological diffusion, absorption or adoption.

[3]See http://www2.census.gov/econ/susb/data/us_state_totals_2007-2012.xls, read 5/21/2014.

[4]Appendix A briefly summarizes the history of large-scale and mass production since the Industrial Revolution, situates Adam Smith (1776) in this regard, and provides some statistics on the current size of the mass scale sector.

goods is already inefficient. Second, market power, expressed as a mark up, implies that costs appear relatively lower and so there are lower incentives to innovate. Even so, it may be that some market power is optimal. Aghion et al (2001, 2005) investigate this question, asking how much market power is optimal for innovation, and argue that innovation is efficient at intermediate level of competition. We discuss how to implement such a public policy for reducing market power to optimal levels of competition (see the market power tax below).

By combining different kinds of competition in one model, we open a new pathway for thinking about and testing the efficiency and equity properties of specific market settings.[5] We also suggests public policies that can improve the performance of the mass market economy, in both economic growth and income distribution, complementing free market policies. This analysis is important not only for developed countries, but for the global economy, which approaches a purer form of market functioning than individual countries, since public policies are much weaker at the global level. It is also important for underdeveloped countries, which can also be characterized in terms of these two sectors (Mayer-Foulkes, 2015a, 2015b).

Schumpeterian theory explains how the innovating sector causes economic growth. It analyzes the role of competition and market structure in optimizing this growth; firm dynamics; development and appropriate growth institutions; and long-term technological waves (Aghion, Akcigitz and Howitt, 2013). However, a complete Schumpeterian analysis requires considering not only the innovating sector, but also the small scale, competitive sector, that does not innovate but instead absorbs technologies. The interaction between the two sectors defines the growth rate, the wage level, profits, and overall efficiency.

The interaction between technological change and labor earnings and their distribution has been extensively studied in the context of labor markets and their changing institutions (e.g. Katz and Autor, 1999; Gordon and Dew-Becker 2008). The determinants of labor earnings include human capital or skills, and the further impact on wages of evolving technologies, skill biased demand shifts and shifting trading opportunities (Acemoglu and Autor, 2011). The different evolution of wages at the top and bottom of the income distribution (Autor, Katz and Kearney, 2006) may be associated with a different evolution of the innovative and competitive sectors. Here we are concerned with the general level of labor earnings across the labor market, and how this is defined by the combination of economy-wide innovation and absorption.

The prevalence of poverty and poverty belts in a context of high technological levels requires explanation as much as relatively low wages. Suffice it to mention that in 2010 there were 14 states (plus the District of Columbia) in the US where 30 percent or more of

[5] For example when two countries trade, it may be that only the large scale sectors export.

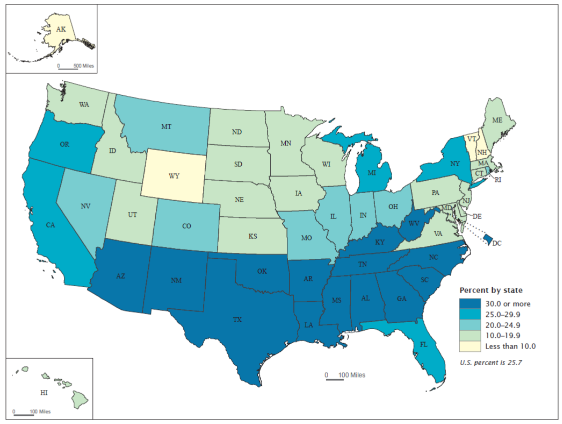


Figure 1. People living in poverty areas by state: 2010. Source: US Census Bureau, 2008-2012 5-year American Community Survey.

the population lived in poverty areas (see Figure 1; Bishaw, 2014).[6] In 2013, 14.5% of the population of the US was poor, 2.0% more than in 2007 (DeNavas-Walt and Proctor, 2014).

The market economy often does not absorb enough of the population into good jobs, both in developed and in underdeveloped countries. In the context of full employment models, what this means is that wages are low compared to the potential technological level. Yet, wages are predicted to be proportional to the technological level both in competitive general equilibrium models, and in models of endogenous technological change. Nevertheless, whole sectors of the population subsist in a series of small scale activities with relatively low technological levels. The empirical section shows that our key indicators of large scale production are causal factors of poverty rates. In this context it is pertinent to discuss the impact of creative destruction on the labor market.

Inequality is not a new phenomenon in the market economy. Piketty (2014) shows that the concentration of income and wealth has been a prominent economic feature since the early 19th Century. However, Piketty mainly presents data. His main analytic tool, the size of the gap between the rate of return on capital r and the economy's growth rate g, which focusses mainly on the financial system, has a limited scope (Piketty, 2015). The model presented here explains the concentration of income and wealth in real terms, not just through the financial system. The cointegration tests show that the key indicators of large scale production are causal factors of the top one and ten percent income ratios in the US.

[6]Poverty areas are defined as census tracts in which more than 20 percent of the people live below the federal poverty line. Besides the District of Columbia, the 14 states were Alabama, Arizona, Arkansas, Georgia, Kentucky, Louisiana, Mississippi, New Mexico, North Carolina, Oklahoma, South Carolina, Tennessee, Texas and West Virginia.

Another contribution of the paper is to introduce technological change in the small scale sector, extending profit driven Schumpeterian models of technological change to small scale producers that absorb technologies just to keep abreast of competition.[7]

We characterize the small scale sector as including goods for which innovation cannot be financed by obtaining sufficient profit margins over a significant proportion of their market. This could be either for technological reasons, or because improvements cannot be appropriated. This setting of technological change in the small scale sector implies several sources of inefficiency. First, the technological absorption that is conducted is repeated by all producers, and restricted to a small scale effort. Second, these efforts are not pooled to produce better results. Third, unexcludable innovation is not pursued, including the use of mass production techniques when feasible. We refer to these sources of inefficiency as the *public good nature of technological absorption*. They imply that public policies can be applied to raise the productivity of technological absorption in the small scale sector.

Summarizing, the aggregate product of the mass market economy is a function of the technological levels of both its sectors. While the innovative sector leads economic growth, it generates inequality and inefficiency in both production and innovation. At the same time, the small scale sector absorbs technologies inefficiently. These inefficiencies in production, innovation and absorption explain how wages can lag behind their potential level, since the overall level of wages is a positive function of the technological levels of both sectors and a negative function of market power.[8] The dynamic properties of innovation and absorption thus determine some of the static properties of the economy, relatively low wages in particular. Finally, income concentration is explained by the concentration of mass production and innovation profits in very few hands.

Two public economic policies can improve on free market policies for the mass market economy, simultaneously promoting both equity and productivity. The first promotes both industrial and small scale technologies by supporting innovation using taxes on profits, and diffusion using these or other taxes. In the US, the high taxes on profits and the support for science and human capital formation applied during the Great Prosperity, which served to promote both innovation and diffusion, corresponds to this combination. These policies are of course independent of Keynesian macroeconomic policies.

[7]In addition, all agents in the model are myopic decision makers with perfect foresight as their time horizon Δt tends to zero. This is both more realistic (there *is* no perfect foresight!) and simpler. It eliminates the need for a second set of variables for the shadow prices of all goods, and the need to predict all prices and levels forever.

[8]For the purposes of this paper, we assume that there is a single, perfect labor market including skilled and unskilled labor. This means that there is a single expected wage level that takes human capital investment costs into account. We abstract from adjustments to the demand and supply of specific skills. These can generate both wage differentials and the creative destruction of employment across skills.

The second addresses the essential source of inefficiency and inequity in the mass market economy: the incentives that producers have to underproduce so as to make higher profits. It is customary for a model proving inefficiencies to propose a tax than can restore efficiency. We define such a *market power tax*. By taxing profit *rates* above some determined level, this tax provides incentives for producers not to underproduce too much. The market power tax generates incentives that reward *production* rather than profit rates. Moreover, its equilibrium taxation revenue is zero. The result is higher efficiency in production and innovation, and higher equity. The market power tax also makes it possible to reduce market power to an optimal level if necessary, as mentioned above.

Neither a general competitive equilibrium model nor a standard model of endogenous technological change include the full set of features we have mentioned. Yet these features are necessary ingredients for understanding the macroeconomic functioning of the industrial, or mass production, market economy, giving rise to a series of urgent issues such as pro-poor growth, global income concentration, the increased political influence of large corporations under deregulation, sustainability in the face of both poverty and corporate power, the global economic business cycle, and so on.

The article proceeds as follows. In the next section we construct the two sector static economy. Then we define the market power tax, which can be used to increase efficiency and equity. The next section introduces technological change in both sectors, obtains the steady state, and shows both innovation and absorption are inefficient. The following section presents the cointegration analysis for the 50 US states plus DC over the period 1997 to 2011, including a discussion of nonstationary panels and cointegration, a description of the data, and an exposition of tests and results. Finally we conclude.

2. The mass market economy

We define a *mass market economy* consisting of two sectors. The heart of the mass market economy is an industrial, technological, mass production sector characterized by ongoing innovation. Innovation is motivated by the acquisition of market power and generates firms spanning important portions of their markets[9]. These firms generate income concentration. Now, for various reasons innovation cannot be financed for every type of good by obtaining sufficient profit margins over a significant proportion of its market. An important proportion of the working population is employed in a second sector, consisting of many small firms that do not innovate significantly and operate competitively. These small, non-innovating firms (including self-employment and informal economic activity), improve their productivity by expending effort on absorbing technologies developed by the industrial sector, which functions as its technological leader.

[9] We consider even small, specialized, innovating firms part of the innovation, mass production sector if they produce for an important portion of their market.

Relative to each other, the large and small scale sectors display *opposite characteristics*. While the first is innovative and displays market power, the second absorbs technologies and is competitive. As we shall see, the innovative sector is at the same time physically more productive, employing higher technologies, and economically less efficient, diverting resources from production of innovative goods through high prices. In turn, the small scale sector is physically less productive, employing lower technologies, and economically more efficient, since it is more competitive.

For simplicity, we keep the defining distinction between small scale and large scale sectors exogenous.[10] While it might be attractive for this boundary to depend on wage levels, for example in the context of studying development and underdevelopment, when considering developed countries, as in the present article, the boundary might still for practical purposes depend mainly on exogenous, technological determinants.[11]

Consider an economy with two sectors L and S that produce a continuum of tradeable goods indexed by $\eta \in [0,1]$, where each η refers to a good. Large scale sector goods $\eta \in \Theta_L = [0,\xi)$ use a mass production technology and are therefore modelled with all production concentrated on a single large producer that is able to make a profit, while small scale sector goods $\eta \in \Theta_S = [\xi, 1]$ are produced on the small scale, with constant returns to scale, therefore modelled with infinitely many small, identical, competitive producers. We assume $\xi > 0$ for some sectors to innovate, and $\xi < 1$ since not all sectors are amenable to mass production.[12] In each sector technological change is endogenous, with differences due to the different competition structures. For simplicity we abstract from innovation uncertainty and assume that innovation is symmetric within each sector L and S. Thus we are assuming goods $\eta \in \Theta_j$ in each sector $j \in \{L, S\}$ have the same technological level A_{jt}.

Innovation occurs as follows. In the large scale sector L there is for each good $\eta \in \Theta_L$ a single, infinitely lived innovator who invests in innovation and becomes a monopolist, producing in the presence of a competitive fringe that we assume consists of large scale producers. For simplicity we assume that innovation is cheaper for the producing incumbent than for any other innovator, and she therefore has an innovation advantage.[13] Her monopoly therefore persists indefinitely. By contrast, in the small sector S anybody can innovate, so

[10]Similarly we abstract from horizontal innovation including the appearance of new small or large scale sectors, or of sectors that have their origins in small enterprises that become large.

[11]The main point is that not all sectors can be ammenable to mass production. Examples are service sectors, such as health. So long as individuality is valued and people live in their own houses, have their own relationships, care for their own children, and perhaps pursue an extraeconomic meaning to their life, a fully mass-produced society can hardly be conceived. Nevertheless, the productivity of mass technologies clearly shapes society, for example our rural-urban social habitat, and deeply interacts with identity (e.g. Lunt and Livingstone, 1992, Akerlof and Kranton, 2000).

[12]Thus the mass market model is in the interior of the continuum lying between competitive general equilibrium and endogenous technological growth.

[13]This means that creative destruction does not produe incentives for innovation in the model. While this might change the balance of innovation incentives under market power, we include in our discussion of studies

as to reap the productive benefits of new technologies, namely the availability of returns to production factors, in this model labor.

We assume that small producers can produce any good, while large producers can only produce goods in sector Θ_L for which mass production technologies are available that are more productive than small scale technologies.

2.1. **Production an consumption.**

2.1.1. *Two kinds of producers.* For simplicity we exclude capital from the production function and limit ourselves to innovation as the source of market power. [14] Thus we only distinguish the two sectors by their competitive context.

Definition 1. *The production function for goods $\eta \in \Theta_j$ in sector $j \in \{L, S\}$ is:*

$$y_{jt}(\eta) = A_{jt} l_{jt}(\eta), \quad j \in \{L, S\} \,.\blacksquare \tag{2.1}$$

Here $y_{jt}(\eta)$ represents the quantity produced of good $\eta \in \Theta_j$. A_{jt} is the technological level in each sector. $l_{jt}(\eta)$ is the quantity of labor input. We assume that the small scale sector can produce any kind of good. The large scale sector is, and must always be, ahead in productivity,

$$A_{Lt} > A_{St}. \tag{2.2}$$

2.1.2. *Preferences.* Let the instantaneous consumer utility $U = U(C_t)$ depend on a subutility function C_t for an agent consuming $c_t(\eta)$ units of goods $\eta \in [0, 1]$, according to the Cobb-Douglass function[15]

$$\ln(C_t) = \int_0^1 \ln(c_t(\eta))\, d\eta. \tag{2.3}$$

Suppose a consumer has a budget z_t for purchasing quantities c_t^L, c_t^S of goods produced in the large and small scale sectors. We assume large and small scale sector goods $\eta \in \Theta_L$, Θ_S are symmetric so have common prices p_{Lt}, p_{St}. Since the composite good kernel (2.3) is Cobb Douglass, consumers dedicate the same budget to each good $\eta \in [0, 1]$. This budget is

---

on optimal levels of competition, such as Aghion et al (2001, 2005). This also means that the incumbent can benefit from any idea that anybody has that benefits her sector.

[14] In constructing the model we attempted to use fixed costs and/or increasing returns to scale in addition to innovation, but both gave rise to mathematics that were too complex for the present purpose. It is worth noting that in the case of fixed costs two equilibria arise for the two sector economy developed here, as in Murphy, Shleifer, and Vishny (1989). In this case realized returns to scale weaken the large scale sector's demand for labor, raising small scale sector employment and therefore reducing wages.

[15] This is also the utility function used in Aghion et al (2005).

z_t, so the quantity bought of each type of good is

$$c_t^L = \frac{z_t}{p_{Lt}}, c_t^S = \frac{z_t}{p_{St}}. \tag{2.4}$$

Hence the quantity of composite good produced is given by $\ln C_t = \xi \ln \frac{z_t}{p_{Lt}} + (1-\xi) \ln \frac{z_t}{p_{St}}$, that is, $C_t = \frac{z_t}{p_{Lt}^{\xi} p_{St}^{1-\xi}}$. Given a budget $p_{Lt}^{\xi} p_{St}^{1-\xi}$, the amount of composite good produced is $C_t = 1$. Letting the composite good be the numeraire, this costs 1, so

$$p_{Lt}^{\xi} p_{St}^{1-\xi} = 1. \tag{2.5}$$

2.1.3. *Choice of production quantities.* Let w_t be the domestic wage level, and suppose now that z_t is the constant expenditure level across goods. Note that therefore aggregate net income is $Z_t = \int_0^1 z_t d\eta = z_t$.

In the case of small producers one unit of good $\eta \in \Theta_S$ is produced competitively by infinitely many firms. Wages equal the income from selling the product of one unit of labor, so the price can be written

$$p_{St} = \frac{w_t}{A_{St}}. \tag{2.6}$$

In each sector $\eta \in \Theta_S$ let $l_{St}(\eta)$ be the aggregate employment of all of the firms producing this good. Since the number of units produced is $c_{jt}^S = \frac{z_t}{p_{St}} = A_{St} l_{St}(\eta)$, the labor quantity is constant in η, so we drop η from the notation, and

$$l_{St} = \frac{z_t}{p_{St} A_{St}}. \tag{2.7}$$

In the case of the large scale sector, each producer has two types of potential competitors. The first type of competitors are small-scale producers, who can produce good η using a technological level A_{St}. Hence it will always be necessary that $p_{Lt} \leq p_{St}$, mass production just being feasible at equality. The second type of competitor, in the competitive fringe, has a lower technological level $\chi^{-1} A_{Lt}$, where $\chi > 1$ represents the competitive edge. This competitor produces on a large scale, supplies the full market, and is just unwilling to enter at zero profit. The incumbent will keep to a maximum price level just at the feasibility level for her competitor. We can think that other potential industrial competitors have even lower technologies for the production of this particular good η.

The level of production considered by both the incumbent and her competitor are given by the aggregate expenditure level on this good, $z_t = p_{Lt}(\eta) y_{Lt}(\eta)$. The maximum markup that the incumbent can use will be χ. Unless we are considering a transition for which mass-production comes into existence with low levels of technological advantage, the usual case will be when under the full markup χ nevertheless $p_{Lt} \leq p_{St}$. The incumbent will drive her industrial competitor to the zero profit limit, and therefore act as if her productivity were A_{Lt}/χ. Hence instead of (2.6) we have

$$p_{Lt} = \frac{\chi w_t}{A_{Lt}}. \tag{2.8}$$

The incumbent produces the same quantity but employing less labor,

$$l_{Lt} = \frac{z_t}{p_{Lt} A_{Lt}} = \frac{\chi^{-1} z_t}{w_t} \tag{2.9}$$

therefore at a cost $\chi^{-1} z_t$, hence making a profit

$$\pi_{Lt} = (1 - \chi^{-1}) z_t. \tag{2.10}$$

2.1.4. *Wages and prices.* The wage level can now be obtained by substituting (2.8), (2.6) in (2.5), so that $1 = p_{Lt}^{\xi} p_{St}^{1-\xi} = \left[\frac{\chi w_t}{A_{Lt}}\right]^{\xi} \left[\frac{w_t}{A_{St}}\right]^{1-\xi}$. Hence

$$w_t = \chi^{-\xi} A_{Lt}^{\xi} A_{St}^{1-\xi}. \tag{2.11}$$

Substituting back in (2.6), (2.8) and simplifying, we can solve for the prices, in terms of the current technological levels

$$p_{St} = \left[\frac{\chi^{-1} A_{Lt}}{A_{St}}\right]^{\xi}, \quad p_{Lt} = \left[\frac{A_{St}}{\chi^{-1} A_{Lt}}\right]^{1-\xi}. \tag{2.12}$$

Hence for large-scale production to outcompete small-scale production at a mark up level χ, the technological levels must satisfy

$$\frac{p_{St}}{p_{Lt}} = \frac{\chi^{-1} A_{Lt}}{A_{St}} > 1. \tag{2.13}$$

We will keep to the case where (2.13) holds, not just (2.2).

2.2. **Labor and income.** Let the population of the economy be $\mathcal{L}$. Suppose $\mathcal{L}_L$ and $\mathcal{L}_S$ are the aggregate employment levels in sectors L and S, with $\mathcal{L}_L + \mathcal{L}_S = \mathcal{L}$. Then if employment levels for each good are l_{Lt}, l_{St}, when the labor market clears,

$$\xi l_{Lt} = \mathcal{L}_L, \quad (1-\xi)\, l_{St} = \mathcal{L}_S, \quad \xi l_{Lt} + (1-\xi)\, l_{St} = \mathcal{L}. \tag{2.14}$$

Now $w_t l_{St} = z_t$, since the participation of labor equals income in sector S, while $w_t l_{Lt} = \chi^{-1} z_t$ in sector L. It follows that

$$\frac{l_{St}}{l_{Lt}} = \chi, \tag{2.15}$$

as also follows from (2.7) and (2.9). Hence, we can solve

$$l_{St} = \frac{\mathcal{L}}{\chi^{-1}\xi + (1-\xi)}, \quad l_{Lt} = \frac{\chi^{-1}\mathcal{L}}{\chi^{-1}\xi + (1-\xi)}. \tag{2.16}$$

From wages and employment income now follows. Using equation (2.11) and (2.16),

$$z_t = w_t l_{St} = \frac{\chi^{-\xi} A_{Lt}^{\xi} A_{St}^{1-\xi} \mathcal{L}}{\chi^{-1}\xi + (1-\xi)}. \tag{2.17}$$

The average wage participation is

$$\frac{w_t \mathcal{L}}{z_t} = \chi^{-1}\xi + (1-\xi). \tag{2.18}$$

Wage participation in the large scale sector is lower than in the small scale sector, so as ξ rises, wage participation drops.

2.3. **Efficiency and equity under market power.** While in some situations there may be a trade-off between efficiency and equity, market power simultaneously results in less efficiency and less equity. This holds at the macroeconomic level in a mass market economy. The static distortions due to the presence of market power are the following.

Theorem 1. *Market power distorts the mass market economy as follows:*

1) Aggregate income is decreasing in market power.

2) The profit to income ratio is increasing in market power.

3) Wages and aggregate wage participation are decreasing in market power.

4) Employment intensity l_{Lt} in the large scale sector is decreasing in market power, while employment intensity l_{St} in the small scale sector is increasing in market power.■

All proofs are in Appendix B.

When the proportion ξ of mass producing sectors increases, the presence of market power implies that wages do not rise in proportion to the increased productivity. Let us examine how the relative size of the large scale sector ξ affect wages in the presence of market power χ.

Theorem 2. *When the size of the large scale sector increases, wages respond as follows:*

$$\frac{\partial \ln w_t}{\partial \xi} = \frac{\partial}{\partial \xi}\left(-\xi \ln \chi + \xi \ln A_{Lt} + (1-\xi) A_{St}\right) = \ln \frac{A_{Lt}}{\chi A_{St}} < \ln \frac{A_{Lt}}{A_{St}}. ■ \tag{2.19}$$

Note that the impact of mass production on wages can be low if market power is near its maximum feasible level $\chi = \frac{A_{Lt}}{A_{St}}$, when the large scale sector faces low large scale competition. Furthermore, when new large scale sectors do not face competition from small scale sectors, so that χ can be larger than $\frac{A_{Lt}}{A_{St}}$, the impact on wages could be negative.

2.4. **Efficiency and inequality in the presence of capital.** The simplest stylized model for a mass market economy, as presented here, does not require the inclusion of capital. However, there are some interesting points that can be made if capital is included. First we show theoretically that the economy wide capital to labor ratio and the corresponding wage level can be considerably distorted by market power in the innovating, mass production sector. Let us suppose that we replace Definition 1 with:

Definition 2. *In the presence of capital the production function for goods* $\eta \in \Theta_j$ *in sector* $j \in \{L, S\}$ *is:*

$$y_{jt}(\eta) = \frac{1}{\varepsilon}\left[k_{jt}(\eta)\right]^{\alpha}\left[A_{jt}l_{jt}(\eta)\right]^{\beta}, \quad j \in \{L, S\}, \tag{2.20}$$

where $\alpha + \beta = 1$, $\varepsilon = \alpha^{\alpha}\beta^{\beta}$.■

Here $k_{jt}(\eta)$ represents units of composite good (2.3) defined for consumption, also used now for capital investment. I assume capital markets are perfect so that the interest rate equals the marginal return of capital. Writing K_t for aggregate capital, it can be shown that the interest rate and wages are given by:

$$r_t = \left(\chi^{-\xi}A_{Lt}^{\xi}A_{St}^{1-\xi}\frac{\alpha\mathcal{L}}{\beta K_t}\right)^{\beta}, \quad w_t = \left(\frac{\beta K_t}{\alpha\mathcal{L}}\right)^{\alpha}\left(\chi^{-\xi}A_{Lt}^{\xi}A_{St}^{1-\xi}\right)^{\beta}. \tag{2.21}$$

These can be verified simply by observing that the large scale sector acts as if its technological level were A_{Lt}/χ, so that the effective average technological level across goods is $\chi^{-\xi}A_{Lt}^{\xi}A_{St}^{1-\xi}$.

Suppose for this discussion that the equilibrium interest rate r^* is determined by intertemporal preferences setting $r^* = \rho$. Then the optimal capital to labor ratio is given by:

$$\frac{K^*}{\mathcal{L}} = \frac{\alpha\chi^{-\xi}A_{Lt}^{\xi}A_{St}^{1-\xi}}{\beta\rho^{\frac{1}{\beta}}}. \tag{2.22}$$

At this level of capital per worker the corresponding wage level is:

$$w^* = \chi^{-\xi}A_{Lt}^{\xi}A_{St}^{1-\xi}\rho^{-\frac{\alpha}{\beta}}. \tag{2.23}$$

Hence we have shown:

Theorem 3. *A market power level* χ *reduces both* $\frac{K^*}{\mathcal{L}}$ *and* w^* *by a factor* $\chi^{-\xi}$.■

The interaction of innovation profits in the large scale sector with the interest rate on capital in the small scale, competitive sector, provides a context for understanding the role of the stock market in bringing forward innovation profits, capitalizing innovation income streams according to the prevailing interest rate, and concentrating them on innovators. In the presence of capital, innovation investment yields a profit rate π_{Lt} that is higher than the interest r_t. A capital market provides innovators with an instrument to bring their profit flow to the present. They can sell through the stock market a project producing an income flow through their innovation. Small investors will purchase this income flow capitalized at a value determined (net of risk) by the interest rate. This brings the innovator's profit flows to the present, included in the project's value. The price of the innovative goods will still reflect the original markup. However, the project's book values will not register innovation profits, only a cost for the purchase of technology that already includes the profit accrued to the innovator. Examples when profits are brought forward are: when a company goes public, when a start up is sold, or when mergers or other reorganizations occur. Hence the study of operating profits through accounting books may not address the full impact of innovation profits.

There is a considerable, controversial, literature on the efficiency costs of monopoly power. In a well known paper, Harberger (1954) concentrates on the misallocation costs of monopoly, and arrives at a very low estimate of $\frac{1}{10}\%$ of GDP. The data is obtained from accounting books for seventy three manufacturing industries for the period 1924-1928. In that paper a benchmark operating profit rate of 10% is considered normal and its efficiency costs are not estimated. The paper concentrates on the allocation impact of adjusting higher or lower profit rates to the 10% level. Because it concentrates on a section of the innovating sector, only estimates the impact of these allocation adjustments, and takes its information from accounting books, this paper does not address the issues we raise here. Cowling and Mueller (1978) weaken Harberger's (1954) assumptions and arrive at social cost estimates of 7 to 13%.

Our model goes quite a long way in explaining the inequality pointed out by Piketty (2014) for mass market economies. The reasons are the following. First, in our model Piketty's interest rate r in fact refers to the profit rate, which is even more easily greater than the growth rate g. Second, the concentration process we describe works in terms of the returns to real investments. It is not only that large financial accounts can get a preferential rate of return. It is also that large real investments can access the profit rate through innovation rather than just the competitive interest rate through capital investment. While discussing the historical aspects of Piketty's (2014) work is beyond the scope of this paper, we would hypothesize that convergence to equilibrium inequality levels or capital to income ratios is faster than posited by Piketty, and responds significantly in a couple rather than in quite a few decades to substantial changes in profit level determinants. Thus, while the two World Wars may have had the most salient (negative) impacts on capital accumulation, other changes such as the rise and fall of the economic framework of the Great Prosperity (including taxes on profits, human capital investment, financial regulation and welfare), or epochal changes in globalization, have also had highly significant impacts. Our cointegration tests on the impact of the relative size of the large scale sector on the 1 and 10% top income ratios in the US provide evidence for this.

3. A market power tax

It is customary for a model proving inefficiencies to propose a tax than can restore efficiency. Our results show that the presence of market power implies an inefficiency in production levels and wages. If incentives can be found for producers not to diminish their production so as to raise prices and profits, aggregate economic efficiency will rise.

Let us suppose that a series of conditions not modelled here imply the social, economic or political convenience of some socially designated positive profit rate for large scale production, which however is lower than can be obtained in an unregulated market. For example, such a profit rate might be the optimal one for innovation, whose efficiency may follow an inverted-U relationship with the profit rate (Aghion et al, 2001, 2005).

We define a market power tax whose incentives are for the producer to decrease her exercise of market power up to the socially designated profit rate. No taxes are levied at equilibrium. Instead the effect is to increase production, improving both efficiency and equity.

Suppose some markup χ is prevalent for large scale producers. For any feasible markup $\varkappa \in [1, \chi]$ profits will be $\pi_{Lt} = (1 - \varkappa^{-1}) z_t$. Note the profit to input rate is $\frac{1-\varkappa^{-1}}{\varkappa^{-1}} = \varkappa - 1$. Let $\tau(\varkappa)$ be the tax schedule

$$\tau(\varkappa) = \begin{cases} \tau_0 (\varkappa - \varkappa_0) + \phi_L^\pi & \varkappa \geq \varkappa_0, \\ \phi_L^\pi & \varkappa < \varkappa_0. \end{cases} \tag{3.1}$$

Besides the constant profit tax rate ϕ_L^π,[16] above the profit rate $\varkappa_0 - 1$, where $\varkappa_0 \in (1, \chi)$, taxes rise with markup rates. The result is that from this point on profits are higher for higher production levels rather than higher gross profits.

Theorem 4. *1) Under a tax schedule $\tau(\varkappa)$, if $\tau_0 > \frac{1}{\varkappa_0(\varkappa_0 - 1)}$, the economy behaves as if market power has lowered to $\chi_0 = \varkappa_0$. In this example the marginal tax on profits as the markup increases at $\varkappa_0$ is less than 1 so long as the profit rate is less than* 61.8%.

2) The economy can approximate the first best for which $\chi = 1$. Define instead tax schedule (3.1) using $\varkappa_0 = 1$. To avoid the tax, large scale production adjusts to a markup $\varkappa^*(\tau_0) = \sqrt{1 + \frac{1}{\tau_0}}$, which also tends to 1 as $\tau_0 \to \infty$.■

4. Technological change

We define the process of endogenous change for the technological levels A_{Lt}, A_{St} in this two sector economy.

4.1. **Innovation in the large scale sector.** As mentioned above, there is in each mass production sector a single, infinitely lived innovator who can produce an innovation for the next period. We consider an innovator with perfect myopic foresight. This means she maximizes profits in the short term Δt by choosing an innovation input flow, and then lets $\Delta t \to 0$. Mayer-Foulkes (2015a) shows that this is equivalent to defining perfect myopic

[16] This constant rate may respond to other reasons for taxation, including all types of public and social goods and equity, which may raise the preference for taxes (Forslid, 2005). However, a more efficient and equitable society has less unsatisfied needs and may therefore need less taxes.

foresight as having perfect knowledge of the current economic variables' time derivatives. The myopic agent uses this knowledge to maximize the current time derivative of her profits.

The effectiveness of innovation investment of the product η entrepreneur has two components. The first is derived from knowledge and is proportional to the skill level $S_{Lt} = A_{Lt}$ that she has been able to accumulate in production, which we assume is the technological level of her firm. The second component is a material input flow v. Innovation occurs with certainty combining these components to obtain a technological level rate of change at time t given by:

$$\frac{\partial}{\partial \Delta t} \tilde{A}_L (t + \Delta t, v)\Big|_{\Delta t = 0} = \mu_L S_{Lt}^{\upsilon} v^{1-\upsilon}, \tag{4.1}$$

where $\mu_L > 0$, $0 < \upsilon < 1$.[17] Here $\tilde{A}_L (t + \Delta t, v)$, where $\Delta t > 0$, is a technology trajectory envisaged by the incumbent over a small time interval into the future, given an expenditure level v on innovation. Note that at $\Delta t = 0$, $\tilde{A}_L (t, v) = A_{Lt}$. The parameter μ_L represents the innovation productivity of the combined inputs.

Let $\phi_L^{\pi}, \phi_L^{\iota} \in (0, 1)$ represent a profit tax and an innovation subsidy, positive or negative proxies for all distortions and policies affecting profits and the incentives to innovate, and define the effective innovativity:

$$\tilde{\mu}_L = \left(\frac{(1-\upsilon)\left(1-\phi_L^{\pi}\right)}{1-\phi_L^{\iota}} \right)^{\varsigma} \mu_L^{1+\varsigma}. \tag{4.2}$$

Proposition 1. *Under perfect myopic foresight, the incumbent sets the rate of change of her technological level A_{Lt} at:*

$$\frac{d}{dt} \ln A_{Lt} = \tilde{\mu}_L \left(\frac{z_t}{\chi A_{Lt}} \right)^{\varsigma} . \blacksquare \tag{4.3}$$

Since z_t depends on both A_{Lt} and A_{St} a relative scale effects is introduced that complicates the dynamics under perfect foresight once technological change in both variables is considered. This aspect is simplified by using continuous myopic foresight, which precludes the need to predict both variables.

Note that innovation is *decreasing* in market power χ, because, as can be seen by following the proof, the higher the market power, the relatively lower costs are compared to profits and therefore the lower the impact of the cost of technological improvement on profits. In other words, the easier it is to make profits, the relatively less worthwhile to spend on cost-saving innovation.

4.2. **Innovation in the small scale sector.** We introduce technological change in the small scale sector. We thus extend Schumpeterian models of technological change, usually

[17] We deal only with an advanced country here and therefore abstract from convergence and divergence.

driven by profits, to small scale producers that absorb technologies just to keep abreast of competition. However, these small firms with limited resources can only apply a limited set of techniques to produce their technological change. The entrepreneur might for example dedicate some of her time to search for new techniques and solutions to adapt to his productive context. Although we exclude human capital from our model, it would be possible to think of an entrepreneur who has or could hire human capital for this purpose. Recall that each small scale sector is characterized by the property that innovation cannot be financed by obtaining sufficient profit margins over a significant proportion of its market. Thus the nature of this sector makes it inviable to establish large research crews using more sophisticated techniques, and excludes from consideration the techniques of large scale or mass production.[18] Productivity therefore lags behind in the small scale sector in the steady state.

Assume that the entrepreneurs running small scale firms can invest a flow of v units of material input to obtain a technological level $\tilde{A}_S(t+\Delta t, v)$ similar to the one we just saw for the large scale sector, given by an innovation function analogous to (4.1),

$$\frac{\partial}{\partial \Delta t}\tilde{A}_S(t+\Delta t, v)\Big|_{\Delta t=0} = \mu_S\left(\frac{A_{Lt}-A_{St}}{A_{Lt}}S_{St}\right)^{v} v^{1-v}. \tag{4.4}$$

Here μ_S is analogous to μ_L, except that it reflects a limited kind of innovation, the kind of innovation that can be carried out on a small rather than large scale $\mu_S < \mu_L$. This is analogous to the distinction between implementation and R&D in Howitt and Mayer-Foulkes (2005), in that in the small scale innovation is unlikely to use an R&D lab, employ scientists, and so on, and is more likely simply to implement technologies created in the large scale sector.[19] S_{St} is the skill level of the firm (entrepreneur, workers and installed productivity), which we consider equal to A_{St}. Here, however, the small scale sector, which in this setting always lag behind the large scale sector, experiences a technological spillover from the large scale technology A_{Lt}, represented by the factor $\frac{A_{Lt}-A_{St}}{A_{Lt}}$.

Recall that the defining characteristic of the small scale sector is that firms cannot obtain sufficient profit margins over a significant proportion of their market. Thus a significant level of market power cannot be achieved, and we assume producers are price takers. However, they cannot be infinitesimally small and still invest in technological absorption. Thus we assume there is some large number of firms N, which represents an approximation to perfect competition. For simplicity all small scale firms are the same size. Therefore their sales are

[18] Franchises may be contexts in which an innovator has devised a way to transform a small scale sector into a large scale sector.

[19] These new technologies may often already be embodied in capital or inputs, although we abstract from these in this simplified model.

$\bar{z}_t = \frac{z_t}{N}$. Let

$$\tilde{\mu}_S = \frac{1}{N}\left(\tfrac{1-\upsilon}{1-\phi_S^{\iota}}\right)^{\varsigma}\mu_S^{1+\varsigma}. \tag{4.5}$$

Note that the effective technological absorptivity $\tilde{\mu}_S$ is decreasing in N.

Proposition 2. *Under perfect myopic foresight, small scale producers set their rate of technological absorption at:*

$$\frac{d}{dt}\ln A_{St} = \tilde{\mu}_S \tfrac{A_{Lt}-A_{St}}{A_{Lt}}\left(\tfrac{A_{Lt}}{A_{St}}\tfrac{z_t}{A_{Lt}}\right)^{\varsigma}.\blacksquare \tag{4.6}$$

4.3. **The steady state.** We now find the steady state growth rate and the steady state relative lag of the small scale sector.

Definition 3. *Define the relative state variable* $a_t = \frac{A_{St}}{A_{Lt}}$.$\blacksquare$

Writing income (2.17) in the form

$$\frac{z_t}{A_{Lt}} = \frac{\chi^{-\xi}a_t^{1-\xi}\mathcal{L}}{\xi\chi^{-1}+(1-\xi)}, \tag{4.7}$$

substitute in (4.3), (4.6) to express the rates of technological change in terms of the relative technological level a_t,

$$\frac{d}{dt}\ln A_{Lt} = \tilde{\mu}_L\left(\tfrac{\chi^{-1-\xi}a_t^{1-\xi}\mathcal{L}}{\xi\chi^{-1}+(1-\xi)}\right)^{\varsigma}, \quad \frac{d}{dt}\ln A_{St} = \tilde{\mu}_S\tfrac{A_{Lt}-A_{St}}{A_{Lt}}\left(\tfrac{\chi^{-\xi}a_t^{-\xi}\mathcal{L}}{\xi\chi^{-1}+(1-\xi)}\right)^{\varsigma}.$$

The dynamics of the relative technological level a_t between small and large scale production can now be written,

$$\frac{d}{dt}\ln a_t = H(a_t) \equiv \left(\tfrac{\chi^{-\xi}\mathcal{L}}{\xi\chi^{-1}+(1-\xi)}\right)^{\varsigma}\left(\tilde{\mu}_S(1-a_t)a_t^{-\varsigma\xi} - \tilde{\mu}_L\chi^{-\varsigma}a_t^{\varsigma(1-\xi)}\right). \tag{4.8}$$

For simplicity we now assume that the small scale sector cannot overtake the large scale competitive fringe, that is, $\tilde{\mu}_S(1-\chi^{-1}) < \tilde{\mu}_L\chi^{-\varsigma}$, so $H(a_t) < 0$ for $a_t > \chi$. This also implies that condition (2.13) is maintained, so that the large scale sector maintains the market power implied by its markup χ.

Theorem 5. *Suppose that the small scale sector cannot overtake the large scale competitive fringe. The relative technological level a_t of the small to the large scale sector has a unique positive steady state $a^* < \chi^{-1}$ with growth rate growth rate $\gamma^* = \frac{d}{dt}\ln A_{Lt}\big|_{a_t=a^*}$ given by*

$$\frac{\tilde{\mu}_L^{1/\varsigma}a^*}{\tilde{\mu}_S^{1/\varsigma}(1-a^*)^{1/\varsigma}} = \chi, \tag{4.9}$$

$$\gamma^* = \left(\tfrac{\chi^{-\xi}\mathcal{L}}{\xi\chi^{-1}+(1-\xi)}\right)^{\varsigma}\tilde{\mu}_S(1-a^*)a^{*(-\varsigma\xi)}. \tag{4.10}$$

The steady state a^* and growth rate γ^* satisfy:

$$
\begin{array}{llll}
\frac{\partial a^*}{\partial \chi} > 0, & \frac{\partial a^*}{\partial \xi} = 0, & \frac{\partial a^*}{\partial \tilde{\mu}_S} > 0, & \frac{\partial a^*}{\partial \tilde{\mu}_L} < 0, \\
\frac{\partial \gamma^*}{\partial \chi} < 0, & \frac{d\gamma^*}{d\xi} > 0, & \frac{\partial \gamma^*}{\partial \tilde{\mu}_S} > 0, & \frac{\partial \gamma^*}{\partial \tilde{\mu}_L} > 0.
\end{array}
\tag{4.11}
$$

Increases in $\tilde{\mu}_S = \frac{1}{N}\left(\frac{1-\upsilon}{1-\phi_S^{\iota}}\right)\mu_S^{1+\varsigma}$ can be obtained by addressing the public good nature of small scale innovation, mentioned above.■

4.4. **Inefficiency of innovation.** Is the private assignment of innovation resources optimal in the mass market economy? We answer this question by examine the innovation incentives for a benevolent government. We show that it is possible to improve income growth by subsidizing innovation, and explain under what conditions this subsidy can be paid for by taxing profits.

In accordance with perfect myopic foresight, let the government maximize $Z_{t+\Delta t}$, deducting expenses in innovation incurred for raising $Z_{t+\Delta t}$. Note that this optimization assumes market exchange takes place in the presence of market power, so the question posed is only seeking a second best. More precisely, at any time t the government maximizes

$$
\max_{\upsilon_{Lt},\upsilon_{St}} \frac{\partial}{\partial \Delta t}\tilde{Z}\left(t+\Delta t,\upsilon_{Lt},\upsilon_{St}\right)\Big|_{\Delta t=0} - \left[\xi \upsilon_{Lt} + (1-\xi)\, N\upsilon_{St}\right]. \tag{4.12}
$$

Here $\tilde{Z}\left(t+\Delta t,\upsilon_{Lt},\upsilon_{St}\right)$, where $\Delta t > 0$, is an income trajectory envisaged by the government over a small time interval into the future, given an expenditure levels υ_L in innovation investment in each large scale sector, and υ_S in innovation investment by each of the N firms in each small scale sector. The maximization is subject to the physical equations for technological change (4.1) and (4.4). Note that the N small firms still repeat innovation in this government maximization.

Now, using expression (2.17), $\frac{d\ln Z_t}{dt} = \xi \frac{d}{dt}\ln A_{Lt} + (1-\xi)\frac{d}{dt}\ln A_{St}$. Hence the government maximization takes the form

$$
\max_{\upsilon_{Lt},\upsilon_{St}} \{\xi\left[\mu_L\left(\frac{\upsilon_{Lt}}{A_{Lt}}\right)^{1-\upsilon} - \frac{\upsilon_{Lt}}{Z_t}\right] + (1-\xi)\left[\mu_S\left(\frac{A_{Lt}-A_{St}}{A_{Lt}}\right)^{\upsilon}\left(\frac{\upsilon_{St}}{A_{St}}\right)^{1-\upsilon} - \frac{N\upsilon_{St}}{Z_t}\right]\}. \tag{4.13}
$$

The first order conditions for (**??**) are:

$$
\frac{(1-\upsilon)\,\mu_L}{A_{Lt}}\left(\frac{\upsilon_{Lt}}{A_{Lt}}\right)^{-\upsilon} = \frac{1}{Z_t}, \quad \frac{(1-\upsilon)\,\mu_S}{A_{St}}\left(\frac{A_{Lt}-A_{St}}{A_{Lt}}\right)^{\upsilon}\left(\frac{\upsilon_{St}}{A_{St}}\right)^{-\upsilon} = \frac{N}{Z_t}. \tag{4.14}
$$

Hence the government would assign innovation expenditures as follows:

$$
\frac{\upsilon_{Lt}}{A_{Lt}} = \left(\mu_L\,(1-\upsilon)\,\frac{Z_t}{A_{Lt}}\right)^{\frac{1}{\upsilon}}, \tag{4.15}
$$

$$
\frac{\upsilon_{St}}{A_{St}} = \left(\mu_S\frac{(1-\upsilon)Z_t}{NA_{St}}\right)^{\frac{1}{\upsilon}}\frac{A_{Lt}-A_{St}}{A_{Lt}}. \tag{4.16}
$$

When these are compared to (9.9) and (9.11), the conditions for obtaining the same resource assignment for innovation are:

$$\frac{1-\phi_L^\pi}{(1-\phi_L^\iota)\chi} = 1, \quad \frac{1}{1-\phi_S^\iota} = 1. \tag{4.17}$$

These equalities are satisfied if $\phi_S^\iota = 0$ and $\frac{1-\phi_L^\pi}{1-\phi_L^\iota} = \chi$. The latter implies $\phi_L^\iota > \phi_L^\pi$.

Thus, except for the N-fold repetition of absorption that occurred in the small scale sector, the fact that these efforts were not pooled, and that non excludable innovation was not pursued, absorption is efficient in the small scale sector. However, large scale innovation is not efficient. The reason was stated above. The easier it is to make profits, the relatively less it is worth to spend on innovation.

The following efficiency results for appropriate government incentives for innovation in the large and small scale sectors can now be stated.

Theorem 6. *1) As market power tends to zero, when $\chi \to 1$, privately assigned innovation tends to efficiency.*

2) When the market power tax is applied, as $\varkappa_0 \to 1$, case 1) is approached in the limit.

3) Suppose that in the large scale sector profits are quantitatively higher than optimal innovation investment. Then taxes and subsidies ϕ_L^π, $\phi_L^\iota \in (0,1)$ exist for which the government's budget is balanced and innovation is optimal. If profits are not that high, a lump sum tax on wages is needed to obtain optimal innovation with a balanced budget.

4) The steady state trajectories of both A_{Lt} and A_{St} lag behind what is economically feasible.

5. A cointegration test of the model

The main implication of our mass market economy model is that the large scale sector positively drives economic growth, productivity, wages, profits and therefore top income participation, albeit with problems in efficiency and distribution. We use macroeconomic data on production by firm size at the state level over the period 1997 to 2011 provided by the Census Bureau and its Statistics of US Business (2014) to test this hypothesis. We concentrate on causality from the large scale sector to the aggregate economy and income inequality, and apply the cointegration methodology. To the best of our knowledge there is no previous study of this impact.

5.1. **Nonstationary Panels and Cointegration.** Macroeconomic indicators over time typically show a clear trend. Unit root tests usually confirm nonstationarity of such data, whereas the error term of the pooled regression equation may or may not be stationary. If the error term is stationary, variables are cointegrated. If not, the estimated relationship is spurious and there are no long-run relationships amongst the variables. Likewise, serial correlation in the error term and feedback from the endogenous variable to the explanatory variables, which usually drive and bias estimators, must be taken into account when estimating long-run relationships.

5.1.1. *Data.* The data set for the 50 US states plus DC over the period 1997 to 2011 come from several data sources. The Census Bureau and its Statistics of US Business (2014) offers for each state time series data on the number of firms, establishments, employment, and annual payroll by employment size of enterprise. Unfortunately, human capital and income are absent in the data; payroll represent effective labor and thus to some extent represents human capital whereas income is available as GDP in millions of current US dollars from the Bureau of Economic Analysis (2014).[20] To deflate, we use the U.S. GDP deflator from the World Development Indicators provided by the World Bank (2014). State level income inequality data are available from Frank (2009). Finally, poverty rates are provided by the US Census Bureau.[21]

The variables used are aggregate state income, represented by its log GDP (y), log employment rate ($l - p$), where p is the logarithm of the state's population, GDP per worker ($y - l$), log average wage rate (w), log payroll rate ($w + l - p$), and log payroll over income ($w + l - y$). While there are several size categories available for firms, we select firms with 500 or more employees to represent the large-scale sector. To represent the proportional role of the large scale sector we use the following variables: log employment in the large scale sector over aggregate employment ($l_L - l$), log payroll in the large scale sector over aggregate payroll ($w_L + l_L - (w + l)$), and log average wage rate in the large scale sector over average wage rate in the aggregate economy ($w_L - w$), referred to below as "large-to-aggregate" employment, payroll and wage ratios respectively. Inequality indicators are $top1$, $top10$ and $poverty$ representing income shares and poverty rates respectively. Note that all of the variables are in principle independent of the state's size.

Table C.2.1 in Appendix C gives descriptive statistics of the variables for all states and large-scale firms and their number of firms and establishments. The means of all variables used for the empirical analysis are close neither to their minimum nor maximum value, which indicates that there is no disproportion. The standard deviations of the variables are relatively large and the values are widely dispersed around the mean. Running a simple regression of each variable regarding time to detect time tendencies, all of the variables have a significant positive trend, as shown in the last column of Table C.2.1 in Appendix C. The District of Columbia was not an outlier in any evident way.

[20]However, our stylized model does not account for human capital dynamics related to technology, and therefore does not explain wage differences between the large and small scale sectors.

[21]Table 21, http://www.census.gov/hhes/www/poverty/data/historical/people.html, read 5/24/2015.

5.1.2. *Panel Unit Root, Cointegration, and Weak-Exogeneity Tests.* Levin, Lin and Chu (2002) (LLC) and Im, Pesaran and Shin (2003) (IPS) propose unit roots tests for pooled balanced data. Suppose that a variable is a function of its lagged value, an autoregressive coefficient, an error term and a deterministic component. The LLC testing procedure assumes that each autoregressive (AR) coefficient is the same for all units while the IPS test procedure allows for heterogeneous autoregressive coefficients. However, both tests require cross-sectional independence and therefore ignore cross-sectional correlation and spillovers between states. Baltagi, Bresson and Pirotte (2007) show that this can seriously bias the size of panel unit root tests. While traditional cross-sectional de-meaning can partly deal with the problem, Pesaran (2007) (PES) lists and discusses newly developed panel unit root tests that attempt to overcome the deficiencies of de-meaning. In particular, PES suggest a simple panel unit root test in the presence of cross-sectional dependencies as an alternative by augmenting cross-sectional averages of lagged levels and first differences of the individual series in the test regression.

If the pooled data is shown to be nonstationary, there are two types of cointegration tests: One can either collect the residuals from pooled regressions to test for stationarity of the error term (residual-based tests) or use the corresponding error correction (EC) terms in EC models to test whether the EC term is significant (error-correction based tests). Residual-based tests such as the (augmented) Dickey-Fuller type tests proposed by Kao (1999), the Phillips and Perron tests by Pedroni (2004) or the Lagrange Multiplier test by McCoskey and Kao (1998) have in common that the residuals from a long-run relationship are tested for stationarity. Error-correction based tests such as Westerlund (2007), however, test whether the EC term in EC models are significant for individual group or full panel models and account for cross-sectional dependence and serial correlation. In particular, the testing procedure of Westerlund (2007) consists of four panel cointegration tests with a null hypothesis of no cointegration for different deterministic components (i.e. zero, constant, trend).

Finally, the validity of the inference of the cointegration vector obtained from EC models in a panel context depends, similar to time series analysis, on the assumption of weak exogeneity. While in time series weak exogeneity tests are conducted for each cross-sectional unit, it is still difficult to adapt individual tests to a panel as whole. According to Moral-Benito and Servin (2014) there are different options in a panel context ranging between testing weak exogeneity on average across all cross-sectional units and joining individual tests into a panel-wide test. In particular, Moral-Benito and Servin (2014)'s panel weak exogeneity test for cointegrated panels is based on the maximum of individual Wald test statistics, where they use the long-run parameter of the cointegration vector and test if they don't show error correction behavior (i.e. are weak exogeneous). Their Monte Carlo simulation show that the proposed testing procedure performs well in commonly used sample sizes–in particular with a large T.

5.1.3. *Estimation Techniques: Dynamic OLS.* The presence of cointegration and unit roots considerably affects the asymptotic distributions in both time series and panel analysis. However, as the number of observations increase in T and N, the OLS estimation of the cointegrated variables converges in the long-run equilibrium to the true value. Nevertheless, for moderate sample size, the estimation bias may remain substantial. Kao and Chiang (2000) show that the OLS estimator has a non-negligible bias in finite samples and that the dynamic OLS estimator performs better in estimating the panel equations than does the OLS estimator with bias correction or the fully modified-OLS estimator. As a result, they propose using the DOLS estimator from Stock and Watson (1993) when dealing with nonstationary panels.

In accounting for endogeneity and serial correlation in the error term, the DOLS estimator considers leads and lags of the first differences:

$$y_{i,t} = \delta_i' d_t + x_{i,t}' b + \sum_{j=-q_1}^{q_2} c_{ij} \Delta x_{i,t+j} + e_{i,t}, \quad i = 1, ..., N \text{ and } t = 1, ..., T, \tag{5.1}$$

where d_t represents the deterministic components (i.e. constant, trend), $x_{i,t}$ is a K-dimensional vector (K being the number of explanatory variables), and $e_{i,t}$ is a stationary error term with zero mean. In the case of $d_t = 0$ there is no deterministic term, whereas in the case of $d_t = 1$ or $d_t = (1, t)'$ there is a constant, or a constant and a trend respectively.

5.2. **Empirical Results.** In estimating the long-run relationships between large-to-aggregate ratios, aggregate indicators and inequality indicators, we use the estimation techniques from Stock and Watson (1993) proposed by Kao and Chiang (2000) in order to deal with nonstationary panel data. Hence, we test first if the pooled data has unit roots and second if serial correlation and cross-sectional dependencies exist. We then test if there is cointegration amongst the integrated variables. In particular, we are interested in the long-run relationships between (1) large-to-aggregate ratios and aggregate indicators, (2) large-to-aggregate ratios and inequality indicators and (3) aggregate indicators and inequality indicators. Finally, we apply the DOLS estimator, either contemporaneous or with one lag in the independent variable, to estimate the signs of the cointegrated relationships we have found and test for weak-exogeneity in the cointegrated panel.

5.2.1. *Nonstationary and Cointegration Testing.* Unit roots test must confirm that the pooled data exhibit unit roots and follow a nonstationary path. We apply first the LLC and IPS mitigating the effects of cross-sectional dependence by de-meaning and second the PES test as an alternative. All test procedures have a null hypothesis of nonstationarity, while the alternative hypothesis for LLC is that all individual series and for IPS and PES is that some individual series are stationary with individual first order autoregressive coefficient. The pooled data is assumed to have a constant and a trend, and tests are implemented with one or two lags in the test regression. Unit root test statistics and p-values in parenthesis from LLC, IPS and PES for I(1) in levels are given in Table C.2.2 in Appendix C and for I(0) in first differences in Table C.2.3 in Appendix C. Accordingly, unit root test statistics in the case of the IPS and PES testing procedures confirm unit roots for one or two lags for almost every variable, as the null hypothesis of nonstationarity is not rejected at least at a 5% level. Bearing in mind that the IPS and PES testing procedure allow for heterogeneous autoregressive coefficients, we conclude that the variables are I(1). Turning to the first differences next, tests statistics from all three tests reject the null hypothesis of nonstationarity again for almost every variables using one or two lags. Since our variables are I(1) in levels and I(0) in first differences, we conclude that the pooled data has unit roots and follow a nonstationary path around a trend.

Turning to cointegration tests, Westerlund (2007) allows dealing with serial correlation and cross-sectional dependence for different deterministic components (i.e. zero, constant, trend). Hence, we test, first, for no first order autocorrelation by the Woolridge (2002) and second, for cross-sectional independence by the Pesaran (2004) testing procedures respectively. We than use Westerlund's (2007) four different panel cointegration tests with a null hypothesis of no cointegration. All cointegration tests are implemented pairwise with a constant and a trend in the test regression. Serial correlation and cross-sectional dependences test statistics with their p-values in parenthesis from Wooldrige (2002) and Pesaran (2004) are given in Table C.3.1 to C.3.3 in Appendix C, while cointegration test statistics and the robust p-values from Westerlund (2007) are given in parenthesis in Table C.4.1 to C.4.3 in Appendix C respectively. Starting with the null hypothesis of no first-order autocorrelation, serial correlation test statistics from Wooldrige (2002) (see Table C.4 in Appendix C), show overall evidence of first-order autocorrelation for each test combination of the panel data. Moreover, cross-sectional dependence test statistics from Pesaran (2004) reject the null hypothesis of cross-sectional independencies for each test combination of the panel data no matter if fixed or random effects are assumed according to the Hausman (1978) test. Hence, we take into account that there is serial correlation and cross-sectional dependencies in our panel data. Turning to the null hypothesis of "no cointegration", test statistics from Westerlund (2004) confirm mainly cointegration between large-to-aggregate ratios to aggregate indicators and inequality indicators.

5.2.2. *Estimation Results.* Analyzing the empirical results of our cointegrated panel, we assign a number, a sign and a significance level to the cointegrated relationships between (1) large-to-aggregate ratios and aggregate indicators, (2) large-to-aggregate ratios and inequality indicators and (3) aggregate indicators and inequality indicators, which have been tested previously. The number represents the number of significant cointegration test results (out of four different tests). The sign and significance level of the long-run relationships are estimated by two dynamic OLS estimations – one contemporaneous and one with one lag (see Tables C.5.1 to C.5.3 in Appendix C). In the case of a significant result, a sign and its significance level is assigned if both coefficients have the same sign. This is shown in Table 1 for each of the test combination and represented in Figure 2 as a graph[22]. Note that causality by Table 1 runs from the variables listed first, on the left (i.e. denoted as "x") to the variables listed next, on the top (i.e. denoted by "y"). Some of the arrows in Figure 2 are dotted to make the diagram less cluttered. The absence of an entry means that either the null hypothesis of "cointegration" or the null hypothesis of "weak exogeneity" of cointegrated variables was rejected. The results of the weak exogeneity tests are given in Table C.6.1 to C.6.3 in Appendix C.

Overall, the results clearly support the proposition set out by the theoretical model, that the large-scale sector drives aggregate economic growth.

In particular, we first look at the results for the core variables of the model, the large-to-aggregate ratios, and the aggregate economy indicators. The two variables $l_L - l$ and $(w_L + l_L) - (w + l)$, representing the large-to-aggregate employment and payroll ratios, are causally related towards the economy both directly and indirectly. There is a direct impact of the large-to-aggregate employment ratio on wages, as predicted by the model. The indirect impacts work through the negative of the large-to-aggregate wage ratio $-(w_L - w)$. Regarding this variable what is remarkable is the consistency of the signs. All pairwise DOLS estimates for "arrows" entering and leaving the original positive variable are negative (except for the arrow acting on poverty, see Table 1 and Figure 2). This means that once an inward and outward causal impact combine, it acts positively. Such combined impacts raise aggregate employment, payroll, and wages (and diminish poverty). We surmise that as the large-to-aggregate employment rises (so the two variables $(l_L - l)$ and $((w_L + l_L) - (w + l))$ rise) and results in a larger aggregate demand for employment l, the wage ratio $(w_L - w)$ decreases as labor from the small scale sector is transferred to the large scale sector. This can happen because less skilled labor is employed instead of more skilled labor in the large scale sector, thereby reducing the wage ratio, or because transferable labor becomes scarcer. The wage level w is at the bottom of the causal hierarchy, determined by other variables, the result of an equilibrium process as $(w_L - w)$ is. Finally, aggregate productivity $y - l$ was introduced because it often appears in studies on economic growth. It is positively impacted by the large-to-aggregate employment ratio $(l_L - l)$. To sum up, the causality results in this set

[22] Figure C.1 contains the same results for the full set of pairwise interactions before taking into account the weak exogeneity test.

of variables confirm the hypothesis that the large-to-aggregate ratios drive the aggregate indicators, including aggregate productivity.

Let us turn now turn to the inequality variables $top1$, $top10$, shown in Table 1 and Figure 2. First, large-to-aggregate ratios have a strong positive impact on both income share indicators (using again the negative of the large-to-aggregate wage ratio). Top income participations are increasing in large-to-aggregate employment and payroll ratios but decreasing in large scale wages. Second, aggregate employment and aggregate income per capita are causally related at least to $top1$ income share. Hence, the higher the aggregate payroll and employment, the higher the income inequality.

Finally, poverty rates are positively caused by large-to-aggregate employment and negatively caused by aggregate payroll and employment.

Summarizing, there is a causality structure that runs from the large-to-aggregate ratios to the aggregate economy variables and to the inequality variables. If we restrict our attention to the variables $l_L - l$, $(w_L + l_L) - (w + l)$, $w - w_L$, $w + l - p$, $l - p$, $y - l$, w, $top1$, $top10$ and $poverty$, the matrix representation of the causal interrelations between these variables is lower triangular, so long as large-to-aggregate employment and payroll are considered at the same level of causality, and also aggregate payroll and amployment. The wage ratio $w_L - w$ works as an equilibrium variable intermediating between these two sets of variables. The top income participations are interrelated with each other and with the wage ratio $w_L - w$, are increasing in all economic activity indicators, and define w together with all of the large-to-aggregate ratios and aggregate economy variables. Finally, the poverty rate decreases with payroll and employment, but increases with the large-to-aggregate employment ratio, perhaps through job destruction (see for example Cheremukhin, 2011).

6. Conclusions

Large scale production has been a feature of the market economy since the Industrial Revolution, and developed into mass production since the Second Industrial Revolution, when innovation became a systematic endeavor based on science. The large scale sector, engaged in mass production and innovation, wields market power, and at the same time the small scale competitive sector absorbs the continual flow of new technologies, nevertheless lagging behind.

The mass market model we construct putting together these economic features shows that the large scale sector generates inefficiencies in production and innovation, as well as inequities, due to the impact of its market power. The small scale sector also generates inefficiencies, because technological absorption has public good features that slow it down. Efforts in technological absorption are repeated by producers, and restricted to small scale pursuits, that are not pooled to produce better results. In addition, unexcludable innovation is not conducted. The model shows that market power also reduces steady state capital accumulation and wage levels even with respect to achieved technological levels, let alone their potential levels.

The model implies a causal structure in economic growth running from the relative size of the large scale sector to aggregate economic variables. A cointegration analysis of US data at the state level over the period 1997 to 2011, including information on production by employment size, GDP, top 1 and top 10 percent income shares, and poverty rates, gives strong evidence for this hypothesis.

Free market institutions can support both competition and market power. The optimality of free market policies depends on what state of competition is actually promoted by the economy. In the presence of innovation-for-profits and mass production, the theoretical analysis shows that they can lead to income concentration and that the overall level of wages can remain below its potential level. Our causal empirical analysis shows that in fact the large scale sector does determine income concentration, wage levels, and poverty.

The current levels of income concentration, wages and poverty are not the only possible ones. The model shows that *efficiency and equity in production and innovation can be promoted together by reducing market power*—the essence of Adam Smith's democratic insights on competition—*and by recognizing the public good nature of technological absorption in the small scale sector*. We define a market power tax that encourages *production* rather than *profit rates*, can generate a more equitable mass market economy, yet levies zero taxes in equilibrium. Profit taxes can also be used to generate efficiency and equity.

The challenge is to make mass production, the workhorse of modern wealth, equitable and truly responsive to pressing economic needs. The mass market economy model, which basically consists of analyzing the mutual impact of economic sectors with and without market power, can serve as a basis for understanding a series of economic and political issues that are characterized by the contradictory impacts of innovation and competition on welfare and distribution, such as income concentration, poverty, increased corporate political influence under deregulation, sustainability in the face of both poverty and corporate power, the global business cycle, and so on.

7. References

Acemoglu, Daron and David Autor, “Skills, Tasks and Technologies: Implications for Employment and Earnings,” in O. Ashenfelter and D. Card, eds., Handbook of Labor Economics, Vol. 4, Elsevier, June 2011, chapter 12, pp. 1043–1171.

Aghion, Philippe; Akcigitz, Ufuk; Howitt, Peter (2013). “What Do We Learn From Schumpeterian Growth Theory?” NBER Working Paper No. 18824.

Aghion, Philippe; Bloom, Nick; Blundell, Richard Griffith, Rachel; Howitt, Peter (2005). “Competition and Innovation: An Inverted-U Relationship,” The Quarterly Journal of Economics, MIT Press, vol. 120(2), pages 701-728, May.

Aghion, Philippe; Harris, Christopher; Howitt, Peter; Vickers, John (2001). “Competition, Imitation and Growth with Step-by-Step Innovation,” *Review of Economic Studies*, LXVIII, 467– 492.

Aghion, P.; Howitt, P.; Mayer-Foulkes, D. (2005). “The Effect of Financial Development on Convergence: Theory and Evidence”, *Quarterly Journal of Economics*, 120(1) February.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith and Peter Howitt (2005). “Competition and Innovation: An Inverted-U Relationship,” *The Quarterly Journal of Economics*, Vol. 120, No. 2 (May), pp 701-728.

Akerlof, George A; Kranton, Rachel E (2000). “Economics and Identity,” *The Quarterly Journal of Economics*, Vol. 115, No. 3 (Aug), pp. 715-753.

Published by: Oxford University Press

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney, “The Polarization of the U.S. Labor Market,” American Economic Review, May 2006, 96 (2), 189–194.

Baltagi, BH, Bresson, G & Pirotte, A (2007). “Panel unit root tests and spatial dependence,” *Journal of Applied Econometrics*, 22, 339-360.

Benassi, Corrado; Cellini, Roberto; Chirco, Alessandra (2003). Personal Income Distribution and Market Structure German Economic Review 3(3): 327-338, Feb.

Bishaw, Alemayehu (2014). Changes in Areas With Concentrated Poverty: 2000 to 2010, American Community Survey Reports 27, US Census Bureau.

Cheremukhin, A. A. (2011). "Labor matching: putting the pieces together," Federal Reserve Bank of Dallas Research Department Working Paper No. 1102.

Chiang, M-H & Kao, C (2000). “On the estimation and inference of a cointegrated regression in panel data,” Advances in Econometrics 15, 179-222.

Cowling, Keith and Mueller, Dennis C. (1978). “The Social Costs of Monopoly Power,” *The Economic Journal*, Vol. 88, No. 352 (Dec), pp 727-748.

Credit Suisse (2014). *Global Wealth Report 2014*, Research Institute, Credit Suisse.

DeNavas-Walt, Carmen and Proctor, Bernadette D. (2014). *Income and Poverty in the United States: 2013*, Current Population Reports, United States Cencus Bureau, September.

Everett, Griff; Hitchcock, Stepanie H; Middleton, Jane; Timms, Rosemary H (2006). *Samuel Slater - Hero or Traitor?: The Story of an American Millionaire's Youth and Apprenticeship in England*, Maypole Promotions (Milford).

Fitton, RS (1989). *The Arkwrights: spinners of fortune*, Manchester, UK; New York: Manchester University Press; New York, NY, USA.

Forslid, Rikard (2005). *Economic geography and public policy.* Princeton University Press.

Gengenbach, C; Urbain, J-P & Westerlund, J (2009). "Error correction testing in panels with global stochastic trends," *Research Memoranda* 051, Maastricht: METEOR, Maastricht Research School of Economics of Technology and Organization.

Gordon, Robert J. and Ian Dew-Becker, "Controversies about the Rise of American Inequality: A Survey," NBER Working Papers 13982, National Bureau of Economic Research, Inc May 2008.

Hall, Robert E (1988). "The Relation between Price and Marginal Cost in U.S. Industry," *Journal of Political Economy*, University of Chicago Press, vol. 96(5), pp 921-47, Oct.

Harberger, Arnold C. (1954). "Monopoly and Resource Allocation, *The American Economic Review*, Vol. 44, No. 2, Papers and Proceedings of the Sixty-sixth Annual Meeting of the American Economic Association (May), pp 77-87.

Hart-Davis, Adam (1995). *Richard Arkwright, Cotton King* - On-line: Science and Technology, Issue 2, October 10, www.exnet.com/1995/10/10/science/science.html, accessed 5/21/2015.

Howitt, P. and Mayer-Foulkes, D. (2005). "R&D, Implementation and Stagnation: A Schumpeterian Theory of Convergence Clubs", *Journal of Money, Credit and Banking*, 37(1) Feb.

Im, KS; Pesaran, MH & Shin, Y (2003). "Testing for unit roots in heterogeneous panels," *Journal of Econometics* 115, 53-74.

Katz, Lawrence F. and Autor, David H (1999). "Changes in the wage structure and earnings inequality," in O. Ashenfelter and D. Card, eds., Handbook of Labor Economics, Vol. 3 of Handbook of Labor Economics, Elsevier, June, chapter 26, pp. 1463–1555.

Kao, C & McCoskey, S (1998). "A residual-based test of the null of cointegration in panel data," *Econometric Reviews*, 17, 57-84.

Kao, C (1999). "Spurious regression and residual-based tests for cointegration in panel data," *Journal of Econometrics* 90, 1-40.

Kirkland, Edward C (1961). "Industry Comes of Age, Business, Labor, and Public Policy 1860–1897," in The Economic History of the United States, Volume VI. New York, Rinehart and Winston.

Lamoreaux, Naomi R. (1991). "Bank Mergers in Late Nineteenth-Century New England: The Contingent Nature of Structural Change," *The Journal of Economic History*, 51: 537-557.

Levin, A; Lin, C-F & Chu, C-SJ (2002). "Unit root test in panel data: asymptotic and finite sample properties," *Journal of Econometrics*, 108, 1-24.

Lipton, Martin (2006). "MergerWaves in the 19th, 20th and 21st Centuries," The Davies Lecture, Osgoode Hall Law School, York University, September 14, available at http://osgoode.yorku.ca/.

Lunt, PK ; Livingstone, SM (1992). *Mass consumption and personal identity*, Open University Press, Buckingham, UK.

Marshall, Alfred (1890). Principles of Economics (8th ed.), The Online Library of Liberty, available at http://www.econlib.org/library/Marshall/marP.html, accessed 5/21/2015.

Mayer-Foulkes (2015a). "The Challenge of Market Power under Globalization," *Review of Development Economics*, Vol 19 (2) 244-264.

Mayer-Foulkes (2015b). "Innovation, Absorption and Distribution in Developed and Underdeveloped Industrial Market Economies," mimeo.

Moral-Benito, E. and Serven, L. (2014) "Testing Weak Exogeneity in Cointegrated Panels", Policy Research Working Paper, no. 7045.

Murphy, Kevin M & Shleifer, Andrei & Vishny, Robert W (1989). "Industrialization and the Big Push," *Journal of Political Economy*, University of Chicago Press, vol. 97(5), pp 1003-26, Oct.

Oxfam (2014). *Working for the Few*, 178 Oxfam Briefing Paper, Jan 20.

Pesaran, MH (2004). "General diagnostic tests for cross section dependence in panels," University of Cambridge, Faculty of Economics, *Cambridge Working Papers in Economics* No. 0435.

Pesaran, MH (2007). "A simple panel unit toot test in the presence of cross section dependence," *Journal of Applied Econometrics*, 27, 265-312.

Piketty, Thomas (2014). *Capital in the Twenty-First Century*. Harvard University Press.

Piketty, Thomas (2015). "About Capital in the Twenty-First Century," American Economic Review: Papers & Proceedings, 105(5): 48–53 http://dx.doi.org/10.1257/aer.p20151060.

Rattenbury, Gordon; Lewis, M. J. T. (2004). Merthyr Tydfil Tramroads and their Locomotives. Oxford: Railway & Canal Historical Society. ISBN 0-901461-52-0.

Schumpeter, Joseph A. (1950) Capitalism, Socialism and Democracy (3rd edition). London: Allen and Unwin.

Smil, Vaclav (2005). Creating the Twentieth Century: Technical Innovations of 1867–1914 and Their Lasting Impact. Oxford; New York: Oxford University Press. ISBN 0-19-516874-7.

Smith, Adam (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*, retrieved 5/21/2015 from http://www.ifaarchive.com/pdf/smith_-_an_inquiry_into_the_nature_and_causes_of_the_wealth_of_nations[1].pdf

Stock, J.H., Watson, M.W., 1993. A simple estimator of cointegration vectors in higher order integrated systems, *Econometrica* 61(4), 783-820.

Sutton, John (2007). "Market Structure: Theory and Evidence," *Handbook of Industrial Organization*, Elsevier, edition 1, volume 3.

Tarasov, Alexander (2009). "Income Distribution, Market Structure, and Individual Welfare," available at http://www.iwb.econ.uni-muenchen.de/forschung/veroffentlichungen/ver_tarasov/income_distribution_tarasov.pdf, read 5/27/2015.

U.S. Bureau of Economic Analysis (2014). Regional Data – GDP and Personal Income. http://www.census.gov/econ/susb/historical_data.html.

U.S. Census Bureau (2014). Statistics of US Business – Historical Data. http://www.census.gov/econ/susb/historical_data.html.

UNCTAD (2008). *World Investment Report 2008*. United Nations, New York.

Westerlund, J (2007). "Testing for error correction in panel data," *Oxford Bulletin of Economics and Statistics*, 69, 709-748.

World Bank (2014). World Development Indicators – World Bank National Account Data. http://data.worldbank.org/indicator.

Yarrow, G. K. (1985). "Welfare Losses in Oligopoly and Monopolistic Competition," *The Journal of Industrial Economics*, Vol. 33, No. 4, A Symposium on Oligopoly, Competition and Welfare (Jun.), pp 515-529.

8. Appendix A. Adam Smith and large scale production

When Adam Smith (1776) explained the benefits of competition in a free market, he addressed an economy made mostly of small producers using only labor-powered machines.[23] Britain's first true factory, a water-powered mill, was first built in 1771. Two important patents, 1769 and 1775, were involved in achieving industrial-scale cotton production.[24] Thus Smith formulated his insights on free markets, cast as preferable to monopoly and other rent-seeking policies,[25] *before the Industrial Revolution developed large scale production.* The US Constitution, adopted in 1787, also laid the foundations for democracy and a market economy before the introduction of large scale production in the US in 1790.[26]

Almost a Century later, the Second Industrial Revolution (1867-1914), based on scientific innovation, generated the basic manufacturing sectors such as steel, oil, mining, telephone, automobile (Smil, 2005). These manufacturing sectors, as well as the banking sector, consolidated into huge enterprises in the late 19$^{\text{th}}$ and early 20$^{\text{th}}$ Centuries, in waves of mergers also featuring vertical integration (Lipton, 2006; Lamoreaux, 1991). The Sherman Antitrust Act of 1890 gives a flavor of this era that culminated in mass production with Henry Ford's 1913 assembly line producing a Model T every 93 minutes (Domm, 2009).

Another Century later, mass production remains the basis of modern productivity, and the force behind globalization. As reported by the U.S. Census Bureau, from 1935 to 1992, the average production of the four largest firms in 459 industries was 38.4% of all shipments. Similarly, from 1992 to 2002, the 200 largest manufacturing companies accounted for 40% of manufacturing value added.[27] The world's top 100 non-financial transnational corporations produced 14.1 percent of global output in 2008 (UNCTAD, 2008).

How would Adam Smith have addressed mass production? What is the trade-off between the physical productivity of large scale firms, and the impact of their market power on other economic sectors, on distribution, efficiency and growth? These are the questions we address using today's methodology in Economics.

[23] When he describes production, Smith (1776) mentions machines frequently, but refers neither to engines nor to the use of steam, water or wind power. The first steam locomotive railway was built in 1804 (Rattenbury and Lewis, 2004).

[24] The mill was built by Richard Arkwright at Cromford, Derbishire, and eventually employed more than 800 workers (Fitton, 1989).

[25] "Monopoly of one kind or another, indeed, seems to be the sole engine of the mercantile system" (Smith, 1776).

[26] Samuel Slater brought the secrets of British textile machinery to the US (Everett, 2006).

[27] Data from U.S. Census Bureau – Economic Census. 1992. "Concentration Ratios for the U.S." http://www.census.gov/epcd/www/concentration92-47.xls., read 9/7/2010.

9. Appendix B. Proofs

Proof of Theorem 1. 1) $\frac{d}{d\chi}\left(\chi^{-(1-\xi)}\xi + \chi^{\xi}(1-\xi)\right) = \xi\chi^{\xi-1}\left(1-\chi^{-1}\right)(1-\xi) > 0$, so from (2.17), $\frac{dZ_t}{d\chi} < 0$. 2) See (2.10). 3) See (2.11). 4) See (2.16).■

Proof of Theorem 2. Differentiate (2.11) and note (2.13).■

Proof of Theorem 4. 1) Below $\varkappa_0$, since $\varkappa_0 < \chi$, the incentives are to raise prices to increase profits. Hence firms will select the mark up $\varkappa_0$. Above $\varkappa_0$, the derivative with respect to $\varkappa$ of $(1-\tau(\varkappa))\pi_{Lt}(\varkappa) = (1-\tau_0(\varkappa-\varkappa_0))\left(1-\varkappa^{-1}\right)z_t$ is negative if

$$0 > -\tau_0\left(1-\varkappa^{-1}\right) + (1-\tau_0(\varkappa-\varkappa_0))\varkappa^{-2} \tag{9.1}$$

$$\Leftrightarrow \quad 0 > -\tau_0\left(\varkappa^2-\varkappa\right) + 1 - \tau_0(\varkappa-\varkappa_0). \tag{9.2}$$

For $\varkappa_0 > 1$ to satisfy the inequality we need $\tau_0\left(\varkappa_0^2-\varkappa_0\right) > 1$, that is, $\tau_0 > \frac{1}{\varkappa_0(\varkappa_0-1)}$. The inequality remains valid for $\varkappa > \varkappa_0$ since the next derivative with $\varkappa$, $-\tau_0(2\varkappa-1)-\tau_0 < 0$ for these values. Observe that the marginal tax on profits at $\varkappa_0$ is

$$\left.\frac{\frac{d}{d\varkappa}\left[(1-\tau(\varkappa))\pi_{Lt}(\varkappa)\right]}{\pi_{Lt}(\varkappa)}\right|_{\varkappa=\varkappa_0} = \frac{\varkappa_0^{-2}}{1-\varkappa_0^{-1}} < 1 \tag{9.3}$$

when $\varkappa_0^2 - \varkappa_0 - 1 > 0$, that is, so long as $\varkappa_0 < \frac{1}{2}\sqrt{5} + \frac{1}{2} = 1.618$, which will only stop holding in this stylized case when the profit rate is above 61.8%.

2) Let $\varkappa_0 = 1$. Then the derivative of $(1-\tau(\varkappa))\pi_{Lt}(\varkappa)$ is negative if

$$0 > -\tau_0\left(\varkappa^2-\varkappa\right) + 1 - \tau_0(\varkappa-1) = 1-(1+\varkappa)\tau_0(\varkappa-1), \tag{9.4}$$

that is, for $\varkappa > \varkappa^*(\tau_0) = \sqrt{1+\frac{1}{\tau_0}}$.■

Proof of Proposition 1. The incumbent's mark up, at time $t+\Delta t$ will be $\frac{\chi A_L(t+\Delta t,v)}{A_{Lt+\Delta t}}$. Thus, using myopic perfect foresight, at any given time t she maximizes her expected rate of change of profit

$$\max_v \left[\left.\frac{d}{d\Delta t}\left[\left(1-\phi_L^\pi\right)\left(1-\left(\frac{\chi A_L(t+\Delta t,v)}{A_{Lt+\Delta t}}\right)^{-1}\right)z_{t+\Delta t}\right]\right|_{\Delta t=0} - \left(1-\phi_j\right)v\right]. \tag{9.5}$$

where $\phi_L^\pi, \phi_L^\iota \in (0,1)$ are the profit tax and innovation subsidy.

The first order condition is:

$$0 = \frac{\partial}{\partial v}\left[\left.\frac{d}{d\Delta t}\left[\left(1-\phi_L^\pi\right)\left(1-\left(\frac{\chi A_L(t+\Delta t,v)}{A_{Lt+\Delta t}}\right)^{-1}\right)z_t\right]\right|_{\Delta t=0} - \left(1-\phi_j\right)v\right] \tag{9.6}$$

$$= \left(1-\phi_L^\pi\right)\left(\frac{\chi A_L(t,v)}{A_{Lt}}\right)^{-2}\frac{\chi\frac{\partial}{\partial v}\frac{d}{d\Delta t}A_{Lt}(v)}{A_{Lt}}z_t - \left(1-\phi_j\right), \tag{9.7}$$

since all other terms are zero. Note that since $\tilde{A}_L(t,v) = A_{Lt}$, $\frac{\partial}{\partial v}\tilde{A}_L(t,v) = 0$. Substituting (4.1) and simplifying,

$$0 = (1-\phi_L^{\pi})(1-v)\mu_L S_{Lt}^{v} v^{-v}\frac{z_t}{\chi A_{Lt}} - \left(1-\phi_j\right). \tag{9.8}$$

Letting $\hat{\mu}_L = \frac{(1-v)\left(1-\phi_L^{\pi}\right)}{\left(1-\phi_L^{\iota}\right)}\mu_L$, material inputs v are given by:

$$v_L = \left(\frac{\hat{\mu}_L}{\chi}\frac{z_t}{A_L}\right)^{\frac{1}{v}} S_{Lt}, \tag{9.9}$$

where we add a subscript L for reference. Substituting this result in (4.1), and writing $\varsigma = \frac{1-v}{v}$,

$$\frac{\partial}{\partial \Delta t}\tilde{A}_L(t+\Delta t, v)\Big|_{\Delta t=0} = \mu_L S_{Lt}^{v}\left(\left(\frac{\hat{\mu}_L}{\chi}\frac{z_t}{A_L}\right)^{1/v} S_{Lt}\right)^{1-v} = \mu_L\left(\frac{\hat{\mu}_L}{\chi}\frac{z_t}{A_L}\right)^{\frac{1-v}{v}} S_{Lt}. \tag{9.10}$$

Note now that perfect myopic foresight implies $\frac{\partial}{\partial \Delta t}\tilde{A}_L(t+\Delta t, v)\Big|_{\Delta t=0} = \frac{d}{dt}A_{Lt}$. Substituting $S_{Lt} = A_{Lt}$ and setting $\tilde{\mu}_L = \mu_L\hat{\mu}_L^{\varsigma}$ according to (4.2), (4.3) is obtained.■

Proof of Proposition 2. Small scale innovation (4.4) is now analogous to large scale innovation (4.1) except that μ_L becomes $\mu_S\left(\frac{A_{Lt}-A_{St}}{A_{Lt}}\right)^{v}$, z_t becomes $\bar{z}_t = \frac{z_t}{N}$, and we consider an innovation subsidy $\phi_S^{\iota} \in (0,1)$, but not a profit tax ϕ_S^{π}. Hence the same derivation yields, after simplification, material inputs given by:

$$v_S = \left(\hat{\mu}_S\frac{z_t}{N}\right)^{\frac{1}{v}}\left(\frac{A_{Lt}-A_{St}}{A_{Lt}}\right)A_{St}^{-\varsigma}. \tag{9.11}$$

and rate of technological change given by (4.6).■

Proof of Theorem 5. Since $\lim_{a_t\to 0} H(a_t) = \infty$, $H' < 0$ and $H(1) < 0$ there is a unique steady state $a^* \in (0,1)$ given by $H(a^*) = 0$. Moreover $a^* < \chi^{-1}$ since $H(\chi^{-1}) < 0$. The steady state level a^* is given by (4.9). Since the RHS is increasing, $\frac{\partial a^*}{\partial \chi} > 0$. The growth rate is given by

$$\gamma^* = \frac{d}{dt}\ln A_{Lt}\Big|_{a_t=a^*} = \frac{d}{dt}\ln A_{St}\Big|_{a_t=a^*}, \tag{9.12}$$

which simplifies to (4.10). Now, this expression is decreasing in a^*, and also decreasing in χ, because $\frac{d}{d\chi}\frac{\chi^{-\xi}}{\xi\chi^{-1}+(1-\xi)} = -\frac{\xi(1-\xi)(\chi-1)}{\chi^{\xi}(\xi+\chi-\xi\chi)^2} < 0$. Since $\frac{\partial a^*}{\partial \chi} > 0$, it follows that $\frac{\partial \gamma^*}{\partial \chi} < 0$.

Addressing changes in ξ, note from (4.9) that $\frac{\partial a^*}{\partial \xi} = 0$. Then from (4.10)

$$\frac{1}{\varsigma}\frac{d}{d\xi}\ln\gamma^* = -\ln(\chi a^*) + \frac{\chi-1}{\xi+\chi-\xi\chi} > 0, \tag{9.13}$$

since $\chi a^* < 1$ and $\frac{d}{d\xi}\ln\left(\xi\chi^{-1}+(1-\xi)\right) = -\frac{\chi-1}{\xi+\chi-\xi\chi}$.

Next, by (4.9), $\chi^{\varsigma}\tilde{\mu}_S = \frac{\tilde{\mu}_L a^{*\varsigma}}{(1-a^*)}$ so, differentiating by $\tilde{\mu}_S$,

$$\chi^{\varsigma} = \tilde{\mu}_L \frac{d}{da^*}\left(\frac{a^{*\varsigma}}{1-a^*}\right)\frac{\partial a^*}{\partial \tilde{\mu}_S} = \tilde{\mu}_L a^{*(\varsigma-1)}\frac{a^* + \varsigma(1-a^*)}{(1-a^*)^2}\frac{\partial a^*}{\partial \tilde{\mu}_S}. \tag{9.14}$$

Hence $\frac{\partial a^*}{\partial \tilde{\mu}_S} > 0$. Note that applying (4.9) to (4.10),

$$\gamma^* = \left(\frac{\chi^{-\xi}\mathcal{L}}{\xi + (1-\xi)\chi}\right)^{\varsigma}\tilde{\mu}_L a^{*\varsigma(1-\xi)}, \tag{9.15}$$

so it follows that $\frac{\partial \gamma^*}{\partial \tilde{\mu}_S} > 0$. Similarly $\frac{\chi^{\varsigma}}{\tilde{\mu}_L} = \frac{a^{*\varsigma}}{\tilde{\mu}_S(1-a^*)}$ so, differentiating by $\tilde{\mu}_L$, $\frac{\partial a^*}{\partial \tilde{\mu}_L} < 0$. Differentiating (4.10) with $\tilde{\mu}_L$, $\frac{\partial \gamma^*}{\partial \tilde{\mu}_L} = \frac{\partial \gamma^*}{\partial a^*}\frac{\partial a^*}{\partial \tilde{\mu}_L} > 0$.■

Proof of Theorem 6. 1) When $\phi_L^{\pi} = \phi_L^{\iota} = 0$ and $\chi \to 1$, $\frac{1-\phi_L^{\pi}}{(1-\phi_L^{\iota})\chi} \to 1$ so innovation tends to efficiency.

2) When the incentives of a market power tax hold, χ is replaced by $\varkappa_0$. Thus in the limit the previous case applies.

3) Observe that the function $\phi_L^{\pi} = f(\phi_L^{\iota}) = 1 - \varphi_L + \chi\phi_L^{\iota}$ (for which $\frac{1-\phi_L^{\pi}}{1-\phi_L^{\iota}} = \chi$) satisfies $f(\frac{\varphi_L - 1}{\varphi_L}) = 0$, $f(1) = 1$ and $f'(\phi_L^{\iota}) = \chi > 1$. The government surplus or deficit in establishing taxes and subsidies ϕ_L^{π}, ϕ_L^{ι} is given by

$$G(\phi_L^{\iota}) = \xi\left[f(\phi_L^{\iota})(1-\chi^{-1})z_t - \phi_L^{\iota}v_{Lt}\right]. \tag{9.16}$$

Let us evaluate this government surplus or deficit at $\phi_L^{\iota} = \frac{\chi-1}{\chi}$ and $\phi_L^{\iota} = 1$. In the first case $\phi_L^{\pi} = 0$, while $\phi_L^{\iota} > 0$, so $G(\frac{\varphi_L-1}{\varphi_L}) < 0$. In the second case

$$G(1) = \xi\left[(1-\chi^{-1})z_t - v_{Lt}\right]. \tag{9.17}$$

Since this quantity, aggregate profits minus optimal innovation costs, is positive by assumption,

$$G'(\phi_L^{\iota}) = \xi\left[\chi(1-\chi^{-1})z_t - v_{Lt}\right] > \xi\left[(1-\chi^{-1})z_t - v_{Lt}\right] \geq 0 \tag{9.18}$$

by the same assumption. Hence by the Intermediate Value Theorem there exists $\phi_L^{\iota} \in (\frac{\chi-1}{\chi}, 1)$ for which the government budget is balanced. At this value ϕ_L^{π}, $\phi_L^{\iota} \in (0,1)$. If instead $G(1) < 0$ a lump sum tax on wages is needed to obtain optimal innovation with a balanced budget.

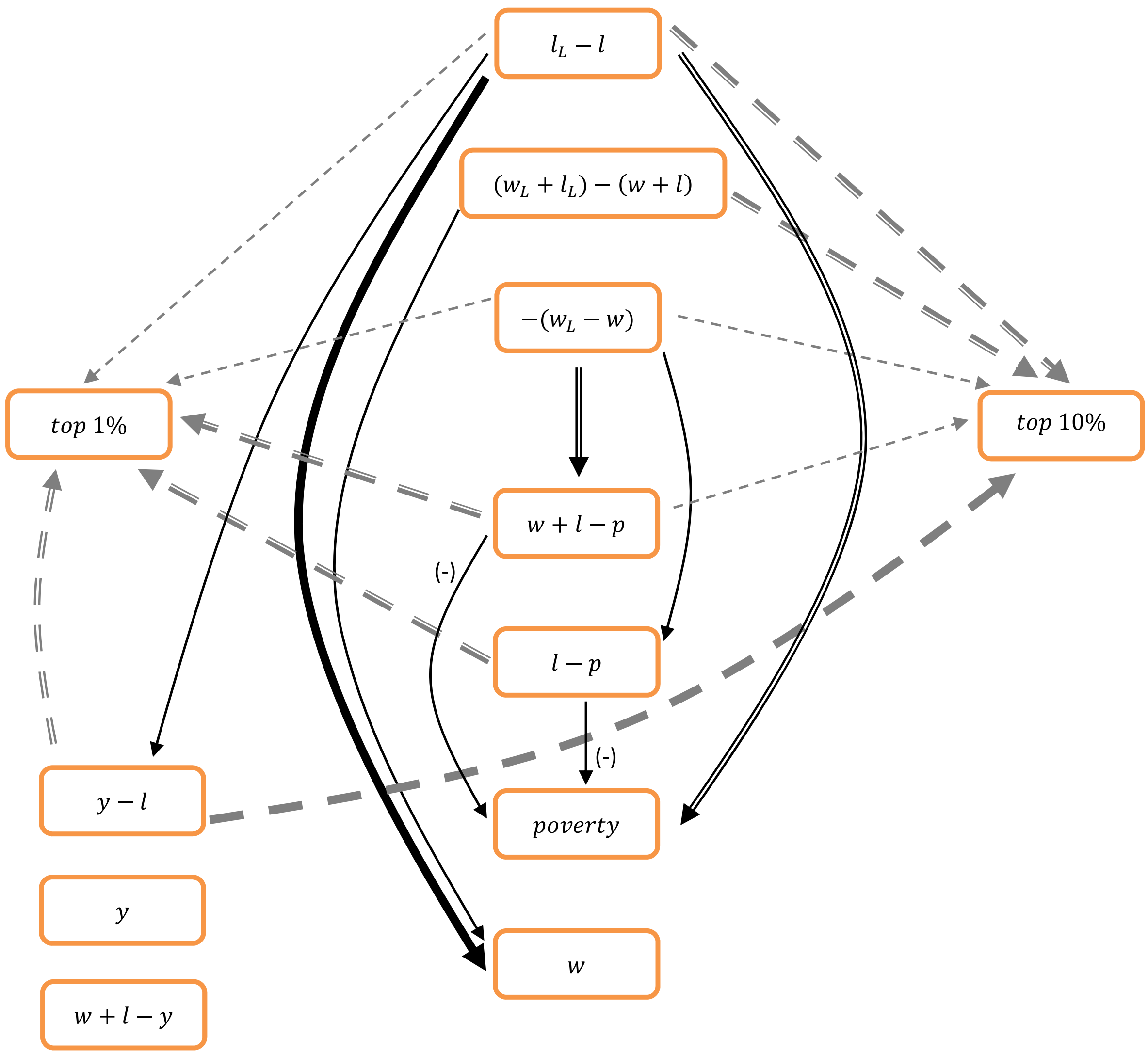
4) The previous statements show this for A_{Lt}. As for A_{St}, in Theorem 5 we showed that by addressing the public good nature of technological absorption and therefore raising $\tilde{\mu}_S$, the small scale sector technological steady state could be raised.■

Table 1: Causality results between large-to-aggregate ratios, aggregate indicators, and inequality indicators (see Tables C.4 to C.6 in the appendix)

x/y	*w + l - y*	*y – l*	*y*	*w*	*l-p*	*w+l-p*	*top10*	*top1*	*poverty*
(w_L+l_L) $-(w+l$				1 (+)***			3 (+)***	2 (+)	
$l_L - l$		1 (+)		4 (+)***			3 (+)***	1 (+)***	2 (+)**
$-(w_L - w)$				3 (+)	1 (+)***	2 (+)***	1 (+)***	1 (+)*	1 (-)
$y - l$							4 (+)***	2 (+)***	1 (–)
l-p								3 (+)***	1 (+)***
$w+l$-p							1 (+)***	3 (+)***	1 (-)***

Notes: The number represents the number of significant cointegration test results (i.e. from a row variable (*x*) to a column variable (*y*)). The absence of an entry means that either the null hypothesis of "cointegration" or the null hypothesis of "weak exogeneity" (in the case of cointegration) was rejected. The sign and significance level of the long-run relationships are estimated by two dynamic OLS estimations – one contemporaneous and one with one lag. In the case of a significant causal relation, a sign and its significance level is assigned if both coefficients have the same sign. The variables l_L and w_L are log employment and wages in firms with 500 or more workers; l and w are log aggregate employment and average wages; y is log aggregate income; w_L+l_L and $w+l$ represent log payrolls; *top1* and *top10* are log income shares of the top 1 and 10% income earners; and *poverty* is the log of the poverty rate.

Figure 2: Causal diagram between large-to-aggregate ratios, aggregate indicators, and inequality indicators (see Tables C.4 to C.6 in the appendix). Each arrow represents significant cointegration and DOLS, and weak exogeneity.

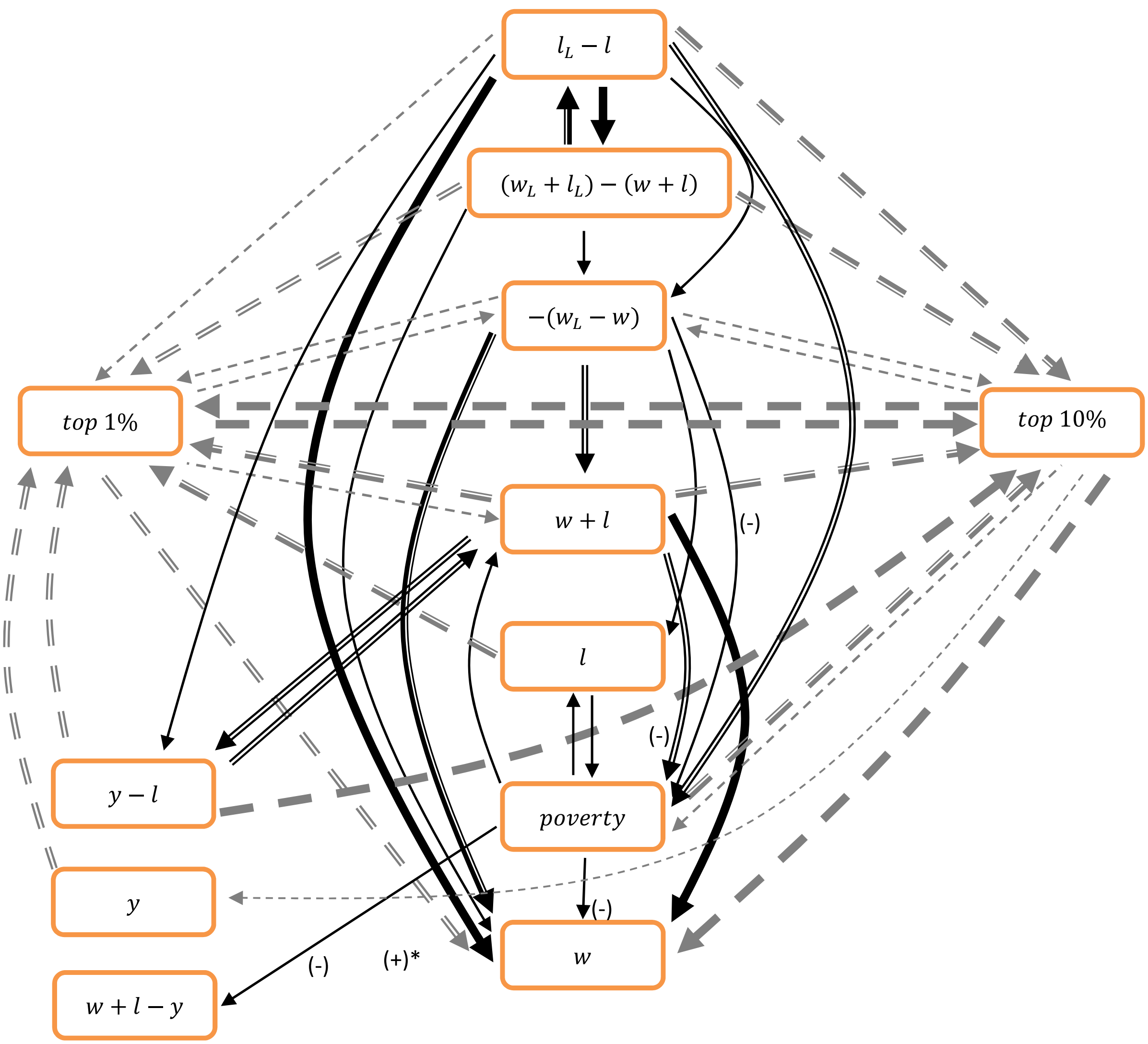


Notes. The types of arrows represent the number of significant entries for each causal relationship: 1 →; 2 ➡; 3 ➡ ; or 4 ➡ . The relationships are positive unless indicated with a negative sign. The sign is assigned as described in the text and in Table 1, and arrows are only represented if they have a consistent significant sign. l_L, w_L are log employment and wages in firms with 500 or more workers. l and w are log aggregate employment and average wages. y is log aggregate income. $top1$ and $top10$ are log income shares of the top 1 and 10% income earners. Some arrows are grey and dotted only to unclutter the diagram. *Poverty* is the log of the poverty rate. $w_L + l_L$, $w + l$ represent log payrolls.

Appendix C. Estimation Figures and Tables

C.1: Cointegration diagram between ratios, economy and inequality indicators

Figure C.1: Cointegration diagram between large-to-aggregate ratios, aggregate indicators, and inequality indicators (see Tables C.4 to C.7 in the appendix)



Note. The types of arrows represent the number of significant entries for each causal relationship: 1 → ; 2 ⇒; 3 ⇛; or 4 ➡ The relationships are positive unless indicated with a negative sign. The sign is assigned as described in the text and in Table 1, and arrows are only represented if they have a consistent sign. l_L, w_L are log employment and wages in firms with 500 or more workers. l and w are log aggregate employment and average wages. y is log aggregate income. $top1$ and $top10$ are log income shares of the top 1 and 10% income earners. Some arrows are grey and dotted only to unclutter the diagram. *Poverty* is the log of the poverty rate. $w_L + l_L$, $w + l$ represent log payrolls.

C.2: Descriptive Analysis and Unit Root Test Results:

Table C.2.1: Descriptive analysis
Annual data for 50 states plus DC from 1997-2011.

		Obs.	Mean	Std. Dev.	Min	Max	Trend
All Firms:							
y	income	765	17.063	1.037	14.975	19.613	0.024 (2.76)***
y-l	income per employee	765	2.936	0.176	2.637	3.634	0.017 (12.59)***
w	average wage rate	765	1.879	0.169	1.514	2.436	0.01 (7.03)***
l	employment	765	14.126	1.029	11.994	16.443	0.007 (0.8)
$w+l$	payroll	765	16.005	1.104	13.638	18.561	0.017 (1.79)*
Number of establishments		765	143,065	152,959	17,680	891,997	915.093 (0.71)
Number of firms		765	116,542	125,261	15,632	730,789	438.749 (0.42)
Firms with 500 or more employees (*L*):							
w_L	average wage rate	765	1.991	0.172	1.622	2.518	0.01 (7.17)***
l_L	employment	765	13.376	1.117	10.81	15.719	0.011 (1.15)
w_L+l_L	payroll	765	15.368	1.184	12.731	17.967	0.021 (2.10)**
Number of establishments		765	20,881	21,764	1,584	120,396	445.296 (2.45)***
Number of firms		765	2,249	1,222	456	5,820	15.254 (1.49)

Notes: Except the number of establishment and the number of firms, all variables in lower letters are used as logarithms and, if measured as a nominal value, are deflated. Large scale firms (*L*) are defined by employment above 500. Running a simple regression of each variables regarding time detects time tendencies. Coefficients with the standard deviation are given parenthesis in the last column.

Table C.2.2: Panel unit root tests with cross-sectional dependence by LLC (2002), IPS (2003) and PES (2007); levels
Annual data for 50 states plus DC from 1997-2011.

	LLC, L(1)	IPS, L(1)	PES, L(1)	LLC, L(2)	IPS, Lag(2)	PES, L(2)
Aggregate Economy						
$(w+l)$-y	-5.989 (0)	-0.9766 (0.16)	0.378 (0.65)	-3.3943 (0)	-1.0617 (0.14)	2.991 (1)
y-l	-5.1105 (0)	0.3 (0.62)	1.577 (0.94)	-2.438 (0.01)	0.3921 (0.65)	2.293 (0.99)
y	-6.8283 (0)	-0.2178 (0.41)	1.231 (0.9)	-3.595 (0)	-0.5819 (0.28)	1.802 (0.96)
w	-8.8018 (0)	-3.0035 (0)	-0.431 (0.33)	-9.8547 (0)	-5.4653 (0)	3.459 (1)
l-p	-6.1139 (0)	0.2949 (0.62)	0.611 (0.73)	-4.6341 (0)	-1.3921 (0.08)	2.037 (0.98)
$w+l$-p	-9.1112 (0)	-2.3248 (0.01)	1.116 (0.87)	-6.4925 (0)	-2.4453 (0.01)	6.661 (1)
Large Scale Ratios						
(w_L+l_L)-$(w+l)$	-7.3068 (0)	-0.9596 (0.17)	-4.618 (0)	-6.8702 (0)	-1.6437 (0.05)	-1.986 (0.02)
l_L-l	-8.7488 (0)	-2.2418 (0.01)	-3.036 (0)	-4.7447 (0)	-1.2613 (0.1)	0.511 (0.7)
w_L-w	-6.5954 (0)	-1.0797 (0.14)	0.756 (0.78)	-4.5714 (0)	-0.5444 (0.29)	2.962 (1)
Income Inequality						
top10	-12.1722 (0)	-4.6892 (0)	-1.281 (0.1)	-16.2029 (0)	-4.6057 (0)	0.065 (0.53)
top1	1.976 (0.98)	4.1766 (1)	-5.705 (0)	11.015 (1)	7.2152 (1)	2.703 (1)
poverty	-7.2445 (0)	-2.9367 (0)	-1.403 (0)	-1.581 (0.06)	-0.864 (0.19)	3.6 (1)

Notes: All variables are used as logarithms (i.e. are given in low letters). The null hypothesis is nonstationarity, while the alternative hypothesis for LLC is that all individual series and for IPS and PES is that some individual series are stationary with individual first order autoregressive coefficient. The adjusted test statistics for LLC (adjusted t*), IPS (*w[t-bar]*) and PES (*z[t-bar]*) and convergence asymptotically to a standard normal distribution. The *p*-values are given in parenthesis. Test statistics account for cross-sectional dependence by removing cross-sectional means from the series in the case of LLC and IPS and by augmenting cross-sectional averages of lagged levels and first differences of the individual series in the case of PES. Tests are implemented with a constant and a trend in the test regression.

Table C.2.3: Panel unit root tests with cross-sectional dependence by LLC (2002) and IPS (2007); first differences

Annual data for 50 states plus DC from 1997-2011.

	LLC, L(1)	IPS, L(1)	PES, L(1)	LLC, L(2)	IPS, Lag(2)	PES, L(2)
Aggregate Economy						
d.((w+l)-y)	-12.965 (0)	-10.9644 (0)	-7.564 (0)	-0.2688 (0.39)	-5.1358 (0)	-1.827 (0.03)
d.y-l	-10.8145 (0)	-9.4298 (0)	-5.6 (0)	-1.763 (0.04)	-5.0788 (0)	-1.637 (0.05)
d.y	-9.8049 (0)	-6.752 (0)	-2.718 (0)	-5.1273 (0)	-4.746 (0)	0.16 (0.56)
d.w	-13.6655 (0)	-10.4273 (0)	-7.138 (0)	-13.5651 (0)	-10.796 (0)	0.563 (0.71)
d.(l-p)	-8.0786 (0)	-6.0768 (0)	-4.124 (0)	-4.3754 (0)	-4.7102 (0)	-0.906 (0.18)
d.(w+l-p)	-11.4319 (0)	-7.6646 (0)	-3.793 (0)	-10.9091 (0)	-6.9474 (0)	2.616 (1)
Large Scale Ratios						
d.((w_L+l_L)-(w+l))	-12.6587 (0)	-9.8821 (0)	-8.754 (0)	-4.9956 (0)	-6.1338 (0)	-4.725 (0)
d.(l_L-l)	-11.9123 (0)	-9.1286 (0)	-6.547 (0)	-5.0654 (0)	-6.0336 (0)	-1.72 (0.04)
d.(w_L-w)	-13.4824 (0)	-11.8791 (0)	-5.078 (0)	-1.2798 (0.1)	-5.7672 (0)	1.601 (0.95)
Income Inequality						
d.top1	-16.5169 (0)	-11.3548 (0)	-4.987 (0)	-13.1909 (0)	-11.109 (0)	-2.089 (0.02)
d.top10	-8.5389 (0)	-6.4092 (0)	-2.23 (0)	3.1442 (1)	-1.2 (0.11)	-0.062 (0.48)
d.poverty	-17.606 (0)	-15.708 (0)	-8.936 (0)	1.565 (0.94)	-6.8503 (0)	0.1 (0.54)

Notes: All variables are used as logarithms (i.e. are given in low letters). The null hypothesis is nonstationarity, while the alternative hypothesis for LLC is that all individual series and for IPS and PES is that some individual series are stationary with individual first order autoregressive coefficient. The adjusted test statistics for LLC (adjusted t*), IPS (*w[t-bar]*) and PES (*z[t-bar]*) and convergence asymptotically to a standard normal distribution. The *p*-values are given in parenthesis. Test statistics account for cross-sectional dependence by removing cross-sectional means from the series in the case of LLC and IPS and by augmenting cross-sectional averages of lagged levels and first differences of the individual series in the case of PES. Tests are implemented with a constant and a trend in the test regression.

C.3: Serial Correlation and Cross-Sectional Dependence Test Results

Table C.3.1: Serial correlation tests by Wooldrige (2002) and cross-sectional dependence tests by Pesaran (2004) in panel-data models; ratio on economy

Annual data for 50 states plus DC from 1997-2011.

Serial Correlation:						
	(w+l)-y	*y-l*	*y*	*w*	*l-p*	*w+l-p*
$(w_L+l_L)-(w+l)$	461.74 (0)	299.3 (0)	612.96 (0)	409.26 (0)	1171.02 (0)	1102.75 (0)
l_L-l	338.61 (0)	326.75 (0)	460.87 (0)	199.52 (0)	988.44 (0)	742.99 (0)
w_L-w	280.2 (0)	346.17 (0)	707.13 (0)	882.96 (0)	763.61 (0)	1629.85 (0)
Cross-Sectional Dependence:						
	(w+l)-y	*y-l*	*y*	*w*	*l-p*	*w+l-p*
Hausman-Test						
$(w_L+l_L)-(w+l)$	0.01 (1)	1.20 (0.55)	-7.75	1.54 (0.46)	0 (1)	0.49 (0.78)
l_L-l	0.24 (0.88)	1.38 (0.5)	-10.38	0.79 (0.67)	5.08 (0.08)	0.45 (0.8)
w_L-w	3.46 (0.18)	0.38 (0.83)	9.75 (0)	0.35 (0.84)	0 (1)	0.13 (0.94)
Pesaran, FE						
$(w_L+l_L)-(w+l)$	58.203 (0)	29.829 (0)	43.399 (0)	52.083 (0)	70.891 (0)	67.55 (0)
l_L-l	52.092 (0)	29.759 (0)	44.088 (0)	46.925 (0)	79.427 (0)	67.411 (0)
w_L-w	57.144 (0)	30.796 (0)	37.974 (0)	54.481 (0)	64.784 (0)	55.702 (0)
Pesaran, RE						
$(w_L+l_L)-(w+l)$	58.191 (0)	29.917 (0)	44.129 (0)	52.032 (0)	70.879 (0)	67.658 (0)
l_L-l	52.248 (0)	29.815 (0)	44.602 (0)	46.766 (0)	78.622 (0)	67.432 (0)
w_L-w	56.68 (0)	30.799 (0)	37.939 (0)	54.412 (0)	64.789 (0)	55.626 (0)

Notes: All variables are used as logarithms (i.e. are given in low letters). Serial Correlation: Tests are implemented pairwise using the residuals from a regression in first differences. The null hypothesis is no first-order autocorrelation in panel data by Wooldrige (2002). F-test statistics and p-values (in parenthesis) are given. Cross-sectional dependencies: Tests are implemented pairwise with a constant and a trend in the test regression using a fixed-effect (FE) or random-effect (RE) model specification. The null hy-pothesis is no systematic difference in coefficients in the case of Hausman (1978), while a cross-sectional independence in panel data is assumed for the null hypothesis in the case of Pesaran (2004). Chi-test statistics and F-test statistics with their p-values (in parenthesis) are given respectively.

Table C.3.2: Serial correlation tests by Wooldrige (2002) and cross-sectional dependence tests by Pesaran (2004) in panel-data models; ratio on inequality

Annual data for 50 states plus DC from 1997-2011.

Serial Correlation:			
	top10	*top1*	*poverty*
$w_L+l_L)-(w+l)$	15.681 (0)	114.159 (0)	77.376 (0)
l_L-l	15.378 (0)	113.946 (0)	77.524 (0)
w_L-w	15.638 (0)	118.360 (0)	78.230 (0)
Cross-Sectional Dependence:			
	top10	top1	poverty
Hausman-Test			
$w_L+l_L)-(w+l)$	9.92 (0)	29.63 (0)	0.64 (0.73)
l_L-l	11.45 (0)	16.19 (0)	5.86 (0.05)
w_L-w	1.46 (0.48)	1.46 (0.47)	3.03 (0.22)
Pesaran, FE			
$w_L+l_L)-(w+l)$	93.511 (0)	82.190 (0)	31.284 (0)
l_L-l	93.049 (0)	84.08 (0)	50.64 (0)
w_L-w	92.775 (0)	85.915 (0)	28.769 (0)
Pesaran, RE			
$w_L+l_L)-(w+l)$	93.472 (0)	85.816 (0)	31.209 (0)
l_L-l	93.196 (0)	86.118 (0)	31.207 (0)
w_L-w	92.825 (0)	86.158 (0)	29.138 (0)

Notes: All variables are used as logarithms (i.e. are given in low letters). Serial Correlation: Tests are implemented pairwise using the residuals from a regression in first differences. The null hypothesis is no first-order autocorrelation in panel data by Wooldrige (2002). F-test statistics and p-values (in parenthesis) are given. Cross-sectional dependencies: Tests are implemented pairwise with a constant and a trend in the test regression using a fixed-effect (FE) or random-effect (RE) model specification. The null hy-pothesis is no systematic difference in coefficients in the case of Hausman (1978), while a cross-sectional independence in panel data is assumed for the null hypothesis in the case of Pesaran (2004). Chi-test statistics and F-test statistics with their p-values (in parenthesis) are given respectively.

Table C.3.3: Serial correlation tests by Wooldrige (2002) and cross-sectional dependence tests by Pesaran (2004) in panel-data models; economy on inequality

Annual data for 50 states plus DC from 1997-2011.

Serial Correlation			
	top10	*top1*	*poverty*
(w+l)-y	17.254 (0)	131.349 (0)	59.577 (0)
y-l	14.450 (0)	116.866 (0)	73.571 (0)
y	10.716 (0)	92.113 (0)	78.701 (0)
w	14.562 (0)	122.377 (0)	75.724 (0)
l-p	10.688 (0)	61.334 (0)	50.541 (0)
w+l-p	9.902 (0)	78.076 (0)	57.057 (0)
Cross-Sectional Dependence			
	top10	*top1*	*poverty*
Hausman-Test			
(w+l)-y	0.14 (0.93)	3.11 (0.21)	2.3 (0.32)
y-l	13.81 (0)	10.14 (0.01)	0.02 (0.99)
y	25.25 (0)	41.49 (0)	105.11 (0)
w	2.18 (0.34)	0.35 (0.84)	30 (0)
l-p	1.96 (0.37)	75.95 (0)	13.56 (0)
w+l-p	0.87 (0.65)	45.28 (0)	41.44 (0)
Pesaran, FE			
(w+l)-y	92.393 (0)	85.989 (0)	18.501 (0)
y-l	92.697 (0)	85.98 (0)	31.794 (0)
y	85.620 (0)	74.072 (0)	18.661 (0)
w	93.766 (0)	86.202 (0)	18.954 (0)
l-p	76.087 (0)	59.476 (0)	14.415 (0)
w+l-p	84.222 (0)	73.309 (0)	7.488 (0)
Pesaran, RE			
(w+l)-y	92.431 (0)	86.232 (0)	19.451 (0)
y-l	93.037 (0)	86.244 (0)	31.74 (0)
y	90.883 (0)	85.486 (0)	29.979 (0)
w	93.832 (0)	86.091 (0)	22.593 (0)
l-p	77.008 (0)	71.840 (0)	15.356 (0)
w+l-p	84.776 (0)	81.274 (0)	11.122 (0)

Notes: All variables are used as logarithms (i.e. are given in low letters). Serial Correlation: Tests are implemented pairwise using the residuals from a regression in first differences. The null hypothesis is no first-order autocorrelation in panel data by Wooldrige (2002). F-test statistics and p-values (in parenthesis) are given. Cross-sectional dependencies: Tests are implemented pairwise with a constant and a trend in the test regression using a fixed-effect (FE) or random-effect (RE) model specification. The null hy-pothesis is no systematic difference in coefficients in the case of Hausman (1978), while a cross-sectional independence in panel data is assumed for the null hypothesis in the case of Pesaran (2004). Chi-test statistics and F-test statistics with their p-values (in parenthesis) are given respectively.

C.4: Panel Cointegration Test Results

Table C.4.1: Panel cointegration tests by Westerlund (2007), ratio on economy; lag (0) and lag(1)
Annual data for 50 states plus DC from 1997-2011.

lag(0))	*(w+l)-y*	*y-l*	*y*	*w*	*l-p*	*w+l-p*
$(w_L+l_L)-(w+l)$						
G_τ	8.375 (0.98)	3.705 (0.65)	0.953 (0.36)	1.093 (0.36)	13.521 (1)	6.253 (0.87)
G_α	9.177 (0.77)	8.772 (0.43)	9.023 (0.63)	7.559 (0.19)	11.083 (1)	9.532 (0.81)
P_τ	7.232 (0.8)	5.733 (0.62)	4.324 (0.38)	2.652 (0.21)	11.11 (0.96)	5.777 (0.65)
P_α	6.322 (0.38)	5.844 (0.18)	6.457 (0.4)	4.764 (0.05)	8.282 (0.92)	6.487 (0.47)
l_L-l						
G_τ	6.937 (0.93)	0.95 (0.4)	2.711 (0.54)	-3.85 (0.08)	15.643 (1)	6.113 (0.85)
G_α	8.116 (0.18)	8.457 (0.24)	9.146 (0.67)	6.966 (0.01)	10.54 (0.94)	9.473 (0.63)
P_τ	5.623 (0.61)	4.616 (0.47)	5.687 (0.58)	-0.707 (0.06)	11.61 (0.96)	6.184 (0.66)
P_α	5.712 (0.14)	5.351 (0.07)	6.409 (0.35)	3.84 (0)	7.431 (0.62)	6.357 (0.3)
w_L-w						
G_τ	3.323 (0.65)	3.08 (0.53)	6.8 (0.92)	-1.798 (0.18)	9.083 (0.95)	0.34 (0.28)
G_α	7.93 (0.12)	8.651 (0.34)	9.741 (0.95)	6.071 (0)	8.582 (0.3)	7.406 (0.03)
P_τ	4.562 (0.53)	6.322 (0.25)	6.377 (0.69)	0.724 (0.11)	7.561 (0.7)	3.529 (0.29)
P_α	5.786 (0.17)	7.153 (0.27)	6.929 (0.58)	4.232 (0.01)	5.777 (0.1)	4.586 (0.02)
lag(1)	*(w+l)-y*	*y-l*	*y*	*w*	*l-p*	*w+l-p*
$(w_L+l_L)-(w+l)$						
G_τ	2.854 (0.64)	6.427 (0.89)	-0.186 (0.3)	-10.49 (0.03)	11.59 (0.99)	8.349 (0.93)
G_α	10.059 (0.41)	10.5 (0.95)	9.972 (0.8)	8.562 (0.78)	11.6 (1)	10.75 (0.93)
P_τ	6.209 (0.55)	7.667 (0.34)	5.41 (0.44)	-3.035 (0)	11.91 (0.97)	7.765 (0.69)
P_α	7.406 (0.3)	7.869 (0.73)	7.827 (0.6)	5.974 (0.17)	8.951 (0.96)	7.923 (0.59)
l_L-l						
G_τ	-3.923 (0.11)	1.01 (0.43)	2.626 (0.57)	-7.932 (0.05)	16.57 (1)	5.103 (0.77)
G_α	9.342 (0.24)	10.42 (0.95)	10.27 (0.86)	8.214 (0.94)	11.05 (0.98)	9.839 (0.44)
P_τ	5.764 (0.28)	8.472 (0.36)	5.593 (0.48)	-2.047 (0)	11.76 (0.95)	3.851 (0.27)
P_α	7.694 (0.49)	7.82 (0.81)	7.954 (0.74)	5.799 (0.42)	8.35 (0.8)	6.924 (0.19)
w_L-w						
G_τ	-1.119 (0.27)	3.123 (0.48)	2.999 (0.62)	-2.506 (0.17)	8.034 (0.86)	1.195 (0.45)
G_α	9.333 (0.17)	10.11 (0.04)	10.41 (0.9)	8.712 (0.92)	10.32 (0.44)	8.885 (0.05)
P_τ	3.678 (0.28)	6.9 (0.01)	7.909 (0.77)	-0.744 (0)	8.898 (0.32)	4.859 (0.41)
P_α	7.057 (0.26)	7.714 (0.06)	7.86 (0.73)	6.239 (0.4)	7.265 (0.07)	6.055 (0.04)

Notes: All variables are used as logarithms (i.e. are given in low letters). Westerlund (2007) presents four different panel cointegration tests with a null hypothesis of no cointegration. While G_τ and G_α test the alternative hypothesis of least one unit is cointegrated (i.e. group mean tests), P_τ and P_α test if the panel is cointegrated as a whole (i.e. panel mean tests). Short run dynamics are restricted to one lag and one lead. The "kernel with" is chosen according to the formula of 4(T/100)^2/9 and therefore to 3. All tests are implemented pairwise with a constant and a trend in the test regression. The robust p-values are given in parenthesis and are based on a bootstrapped distribution using 800 bootstrap replications in order to deal with cross-sectional dependencies.

Table C.4.2: Panel cointegration tests by Westerlund (2007); ratio on inequality, lag(0)
Annual data for 50 states plus DC from 1997-2011.

	top10	*top1*	*poverty*
$(w_L+l_L)-(w+l)$			
G_τ	-5.855 (0.04)	3.409 (0.39)	0.313 (0.28)
G_α	8.814 (0.22)	9.184 (0.36)	8.956 (0.7)
P_τ	-3.98 (0.02)	3.670 (0.08)	1.571 (0.12)
P_α	4.992 (0.04)	5.772 (0.04)	5.956 (0.23)
l_L-l			
G_τ	-1.909 (0.15)	8.309 (0.81)	-4.688 (0.04)
G_α	7.838 (0.05)	9.244 (0.3)	8.311 (0.29)
P_τ	-4.894 (0.01)	5.924 (0.14)	1.939 (0.15)
P_α	4.138 (0.02)	5.674 (0.02)	5.111 (0.06)
w_L-w			
G_τ	3.211 (0.56)	9.833 (0.91)	-0.795 (0.18)
G_α	8.544 (0.2)	9.362 (0.37)	7.866 (0.26)
P_τ	1.682 (0.19)	8.17 (0.42)	1.878 (0.15
P_α	5.472 (0.1)	6.379 (0.08)	5.128 (0.07)

Notes: All variables are used as logarithms (i.e. are given in low letters). Westerlund (2007) presents four different panel cointegration tests with a null hypothesis of no cointegration. While G_τ and G_α test the alternative hypothesis of least one unit is cointegrated (i.e. group mean tests), P_τ and P_α test if the panel is cointegrated as a whole (i.e. panel mean tests). Short run dynamics are restricted to one lag and one lead. The "kernel with" is chosen according to the formula of 4(T/100)^2/9 and therefore to 3. All tests are implemented pairwise with a constant and a trend in the test regression. The robust p-values are given in parenthesis and are based on a bootstrapped distribution using 800 bootstrap replications in order to deal with cross-sectional dependencies.

Table C.4.3: Panel cointegration tests by Westerlund (2007); economy on inequality, lag(0)
Annual data for 50 states plus DC from 1997-2011.

	top10	*top1*	*poverty*
(w+l)-y			
G_τ	4.792 (0.64)	10.56 (0.76)	1.773 (0.43)
G_α	9.592 (0.46)	9.945 (0.24)	8.715 (0.47)
P_τ	3.419 (0.19)	9.274 (0.1)	1.958 (0.18)
P_α	6.388 (0.18)	6.961 (0.04)	5.746 (0.2)
y-l			
G_τ	4.133 (0.05)	12.92 (0.9)	-3.269 (0.09)
G_α	10.44 (0)	10.95 (0.16)	9.298 (0.79)
P_τ	4.105 (0)	11.47 (0.03)	1.508 (0.13)
P_α	7.401 (0)	7.928 (0.03)	6.116 (0.32)
y			
G_τ	2.489 (0.48)	7.485 (0.62)	-0.534 (0.25)
G_α	10.72 (0.96)	11.23 (0.5)	9.442 (0.88)
P_τ	7.793 (0.39)	9.558 (0.02)	1.77 (0.16)
P_α	8.323 (0.49)	8.923 (0.07)	6.929 (0.68)
w			
G_τ	0.095 (0.25)	8.243 (0.81)	-2.67 (0.12)
G_α	9.752 (0.62)	10.64 (0.83)	9.282 (0.79)
P_τ	2.134 (0.16)	9.375 (0.29)	4.918 (0.48)
P_α	6.765 (0.31)	8.155 (0.32)	7.114 (0.73)
l-p			
G_τ	3.758 (0.6)	0.322 (0.13)	-3.214 (0.11)
G_α	9.864 (0.56)	9.66 (0.03)	8.698 (0.59)
P_τ	4.127 (0.47)	2.753 (0)	-0.421 (0.07)
P_α	7.382 (0.56)	7.039 (0.01)	6.522 (0.54)
w+l-p			
G_τ	-5.896 (0.07)	1.708 (0.154)	-4.996 (0.05)
G_α	10.429 (0.8)	9.896 (0.07)	7.679 (0.23)
P_τ	1.417 (0.17)	8.447 (0.02)	1.881 (0.19)
P_α	8.227 (0.87)	7.767 (0.04)	5.941 (0.35)

Notes: All variables are used as logarithms (i.e. are given in low letters). Westerlund (2007) presents four different panel cointegration tests with a null hypothesis of no cointegration. While G_τ and G_α test the alternative hypothesis of least one unit is cointegrated (i.e. group mean tests), P_τ and P_α test if the panel is cointegrated as a whole (i.e. panel mean tests). Short run dynamics are restricted to one lag and one lead. The "kernel with" is chosen according to the formula of 4(T/100)^2/9 and therefore to 3. All tests are implemented pairwise with a constant and a trend in the test regression. The robust p-values are given in parenthesis and are based on a bootstrapped distribution using 800 bootstrap replications in order to deal with cross-sectional dependencies.

C.5: Dynamic OLS Estimation Results

Table C.5.1: Dynamic OLS by Stock and Watson (1993) with lag(1) and lead (1) of first differences; ratio on economy

Annual data for 50 states plus DC from 1997-2011.

	(w+l)-y	*y-l*	*y*	*w*	*l-p*	*w+l-p*
(w_L+l_L)-(w+l)						
L(0)	0.493	0.126	6.478	0.619	0.274	0.893
	*(13.58)****	*(2.23)***	*(24.44)****	*(11.69)****	*(5.75)****	*(10.21)****
L(1)	0.495	0.116	6.414	0.611	0.249	0.859
	*(13.09)****	*(1.97)***	*(23.41)****	*(11.12)****	*(5.06)****	*(9.48)****
l_L-l						
L(0)	0.422	0.042	5.088	0.464	0.289	0.753
	*(13.57)****	*(0.87)*	*(21.18)****	*(9.96)****	*(7.15)****	*(9.99)****
L(1)	0.428	0.031	5.035	0.459	0.271	0.73
	*(13.24)****	*(0.61)*	*(20.40)****	*(9.52)****	*(6.51)****	*(9.36)****
w_L-w						
L(0)	-0.705	0.548	-4.668	-0.158	-0.926	-1.083
	*(-6.47)****	*(3.56)****	*(-4.77)****	*(-0.99)*	*(-7.18)****	*(-4.27)****
L(1)	-0.716	0.548	-4.868	-0.167	-0.884	-1.052
	*(-6.31)****	*(3.41)****	*(-4.82)****	*(-1.02)*	*(-6.62)****	*(-4)****

Notes: All variables are used as logarithms (i.e. are given in low letters). The bias corrected *t*-statistics of the coefficients are reported in parenthesis. * (**) [***] denotes that the coefficient is significantly different from zero at a 10% (5%) [1%] level. All equations include unreported, state-specific constants and a trend.

Table C.5.2: Dynamic OLS by Stock and Watson (1993) with lag(1) and lead (1) of first differences; ratio on inequality

Annual data for 50 states plus DC from 1997-2011.

	top10	*top1*	*poverty*
(w_L+l_L)-(w+l)			
L(0)	0.154	0.096	0.296
	*(5.41)****	*(1.53)*	*(3.17)****
L(1)	0.141	0.036	0.335
	*(4.71)****	*(0.54)*	*(3.51)****
l_L-l			
L(0)	0.15	*0.135*	*0.18*
	*(6.15)****	*(2.49)****	*(2.24)***
L(1)	0.135	*0.066*	*0.204*
	*(5.27)****	*(1.17)*	*(2.49)****
w_L-w			
L(0)	-0.293	-0.316	0.21
	*(-3.67)****	*(-1.77)**	*(0.81)*
L(1)	-0.288	-0.264	0.175
	*(-3.45)****	*(-1.43)*	*(0.66)*

Notes: All variables are used as logarithms (i.e. are given in low letters). The bias corrected *t*-statistics of the coefficients are reported in parenthesis. * (**) [***] denotes that the coefficient is significantly different from zero at a 10% (5%) [1%] level. All equations include unreported, state-specific constants and a trend.

Table C.5.3: Dynamic OLS by Stock and Watson (1993) with lag(1) and lead (1) of first differences; economy on ratio

Annual data for 50 states plus DC from 1997-2011.

	top10	*top1*	*poverty*
(w+l)-y			
L(0)	0.083 *(3.03)******	0.015 *(0.25)*	*-0.627* *(-7.23)******
L(1)	0.065 *(2.25)******	-0.062 *(-1.01)*	*-0.579* *(-6.45)******
y-l			
L(0)	0.25 *(13.98)******	*0.251* *(5.88)******	-0.101 *(-1.55)*
L(1)	0.254 *(13.43)******	*0.247* *(5.46)******	-0.123 *(-1.83)**
y			
L(0)	0.035 *(12.34)******	0.038 *(5.70)******	*0.029* *(2.83)******
L(1)	0.037 *(12.91)******	0.042 *(6.55)******	*0.029* *(2.77)******
w			
L(0)	0.271 *(16.51)******	0.22 *(5.17)******	*-0.394* *(-6.42)******
L(1)	0.264 *(15.59)******	0.182 *(4.17)******	*-0.387* *(6.09)******
l-p			
L(0)	0.195 *(8.78)******	*0.263* *(5.42)******	-0.596 *(-8.28)******
L(1)	0.184 *(7.86)******	*0.188* *(3.86)******	-0.59 *(-7.96)******
w+l-p			
L(0)	0.155 *(15)******	*0.153* *(6.2)******	-0.309 *(-8.31)******
L(1)	0.151 *(14.01)******	*0.122* *(4.95)******	-3.05 *(-7.97)******

Notes: All variables are used as logarithms (i.e. are given in low letters). The bias corrected *t*-statistics of the coefficients are reported in parenthesis. * (**) [***] denotes that the coefficient is significantly different from zero at a 10% (5%) [1%] level. All equations include unreported, state-specific constants and a trend.

C.6: Weak Exogeneity Test Results

Table C.6.1: Weak Exogeneity in Cointegrated Panels by Moral-Benito and Servin (2014); ratio on economy

Annual data for 50 states plus DC from 1997-2011.

	(w+l)-y	*y-l*	*y*	*w*	*l-p*	*w+l-p*
w_L+l_L)-(w+l)	0.944	0.361	0.324	0.285	0.602	0.213
	(0.32)	*(0.5)*	*(0.52)*	*(0.53)*	*(0.42)*	*(0.55)*
l_L-l	-0.758	-1.019	-1.025	-1.055	-0.879	-1.096
	(0.88)	*(0.94)*	*(0.94)*	*(0.94)*	*(0.91)*	*(0.95)*
w_L-w	-2.343	-2.361	-2.353	-2.364	-2.435	-2.367
	(1)	*(1)*	*(1)*	*(1)*	*(1)*	*(1)*

Notes: All variables are used as logarithms (i.e. are given in low letters). The null hypothesis is weak exogeneity of the cointegrated panel (i.e. of all US-states) by Moral-Benito and Servin (2014). The *p*-values are given in parenthesis.

Table C.6.2: Weak Exogeneity in Cointegrated Panels by Moral-Benito and Servin (2014); ratio on inequality

Annual data for 50 states plus DC from 1997-2011.

	top10	*top1*	*poverty*
w_L+l_L)-(w+l)	0.448	0.497	0.492
	(0.47)	*(0.46)*	*(0.46)*
l_L-l	-0.957	-0.952	-0.927
	(0.93)	*(0.93)*	*(0.93)*
w_L-w	*-2.306*	-2.313	*-2.335*
	(1)	*(1)*	*(1)*

Notes: All variables are used as logarithms (i.e. are given in low letters). The null hypothesis is weak exogeneity of the cointegrated panel (i.e. of all US-states) by Moral-Benito and Servin (2014). The *p*-values are given in parenthesis.

Table C.6.3: Weak Exogeneity in Cointegrated Panels by Moral-Benito and Servin (2014); economy on inequality

Annual data for 50 states plus DC from 1997-2011.

	top10	*top1*	*poverty*
(w+l)-y	-2.112	-2.113	-2.146
	(1)	*(1)*	*(1)*
y-l	-0.102	-0.05	-0.4
	(0.67)	*(0.65)*	*(0.78)*
y	*2.612*	2.555	*2.723*
	(0.07)	*(0.08)*	*(0.06)*
w	-1.051	-1.058	-0.977
	(0.943)	*(0.94)*	*(0.93)*
l-p	-1.964	-1.983	-1.949
	(1)	*(1)*	*(1)*
w+l-p	*0.883*	0.83	*0.925*
	0.34)	*(0.35)*	*(0.33)*

Notes: All variables are used as logarithms (i.e. are given in low letters). The null hypothesis is weak exogeneity of the cointegrated panel (i.e. of all US-states) by Moral-Benito and Servin (2014). The *p*-values are given in parenthesis.

C.7: Additional Results of Cointegration and Dynamic OLS

Table C.7.1: Panel cointegration tests by Westerlund (2007); economy vs. economy, lag(0)
Annual data for 50 states plus DC from 1997-2011.

	(w+l)-y	*y-l*	*y*	*w*	*l-p*	*w+l-p*
(w+l)-y						
G_τ	-/-	-5.346 (0.04)	10.35 (0.98)	-5.346 (0.03)	13.92 (1)	7.997 (0.96)
G_α	-/-	8.97 (0.32)	10.49 (0.97)	7.506 (0)	10.84 (0.3)	10.12 (0.96)
P_τ	-/-	0.962 (0.01)	8.163 (0.76)	-0.048 (0)	12.05 (0.12)	10.17 (0)
P_α	-/-	7.113 (0.28)	7.997 (0.86)	4.669 (0)	8.244 (0.19)	7.928 (0.97)
y-l						
G_τ	-3.217 (0.06)	-/-	7.312 (0.92)	-3.216 (0.07)	5.076 (0.73)	1.409 (0.39)
G_α	10.06 (0.92)	-/-	9.982 (0.97)	8.125 (0.11)	10.81 (0.98)	9.881 (0.74)
P_τ	0.16 (0.05)	-/-	7.911 (0.71)	1.272 (0.1)	10.45 (0.65)	4.696 (0.34)
P_α	7.457 (0.46)	-/-	7.591 (0.58)	4.513 (0.02)	8.34 (0.59)	7.007 (0.46)
y						
G_τ	2.379 (0.54)	3.364 (0.66)	-/-	1.374 (0.39)	5.816 (0.81)	-0.348 (0.2)
G_α	10.13 (0.99)	10.28 (0.98)	-/-	9.687 (0.92)	9.66 (0.96)	8.002 (0.47)
P_τ	4.882 (0.44)	6.385 (0.73)	-/-	3.121 (0.27)	8.33 (0.68)	4.639 (0.28)
P_α	7.521 (0.88)	7.75 (0.89)	-/-	6.9 (0.54)	7.637 (0.58)	6.259 (0.25)
w						
G_τ	5.262 (0.83)	5.261 (0.82)	6.225 (0.88)	-/-	5.988 (0.79)	5.988 (0.79)
G_α	9.047 (0.79)	9.988 (0.94)	10.05 (0.96)	-/-	10.24 (0.93)	8.802 (0.54)
P_τ	5.972 (0.68)	5.496 (0.61)	6.383 (0.64)	-/-	9.081 (0.76)	8.003 (0.73)
P_α	7.004 (0.76)	7.365 (0.77)	7.331 (0.61)	-/-	6.795 (0.38)	5.748 (0.15)
l-p						
G_τ	-0.561 (0.28)	0.212 (0.33)	2.819 (0.55)	-2.862 (0.13)	-/-	-2.862 (0.13)
G_α	9.227 (0.84)	9.588 (0.95)	10.17 (0.92)	8.367 (0.27)	-/-	8.071 (0.29)
P_τ	3.383 (0.35)	3.644 (0.38)	7.299 (0.8)	-0.068 (0.1)	-/-	0.158 (0.09)
P_α	6.56 (0.59)	7.134 (0.8)	8 (0.88)	5.356 (0.1)	-/-	5.412 (0.12)
w+l-p						
G_τ	5.841 (0.82)	3.429 (0.61)	6.895 (0.87)	5.735 (0.79)	5.735 (0.77)	-/-
G_α	8.719 (0.92)	9.428 (0.93)	10.22 (0.94)	8.726 (0.42)	9.122 (0.75)	-/-
P_τ	6.213 (0)	5.386 (0.57)	10.82 (0.98)	6.112 (0.53)	4.974 (0.48)	-/-
P_α	6.244 (0.89)	7.214 (0.82)	7.959 (0.89)	5.914 (0.15)	6.633 (0.37)	-/-

Notes: All variables are used as logarithms (i.e. are given in low letters). Westerlund (2007) presents four different panel cointegration tests with a null hypothesis of no cointegration. While G_τ and G_α test the alternative hypothesis of least one unit is cointegrated (i.e. group mean tests), P_τ and P_α test if the panel is cointegrated as a whole (i.e. panel mean tests). Short run dynamics are restricted to one lag and one lead. The "kernel with" is chosen according to the formula of 4(T/100)^2/9 and therefore to 3. All tests are implemented pairwise with a constant and a trend in the test regression. The robust p-values are given in parenthesis and are based on a bootstrapped distribution using 800 bootstrap replications in order to deal with cross-sectional dependencies.

Table C.7.2: Panel cointegration tests by Westerlund (2007); ratio vs. ratio, lag(0)
Annual data for 50 states plus DC from 1997-2011.

	$(w_L+l_L)-(w+l)$	l_L-l	w_L-w
$(w_L+l_L)-(w+l)$			
G_τ	-/-	-5.094 (0.05)	-5.097 (0.05)
G_α	-/-	6.062 (0.02)	8.223 (0.25)
P_τ	-/-	3.106 (0.33)	3.644 (0.32)
P_α	-/-	3.587 (0.02)	6.3 (0.35)
l_L-l			
G_τ	-3.901 (0.07)	-/-	-3.9 (0.08)
G_α	6.963 (0.08)	-/-	8.149 (0.45)
P_τ	0.371 (0.08)	-/-	2.176 (0.18)
P_α	3.776 (0.02)	-/-	5.507 (0.26)
w_L-w			
G_τ	1.596 (0.4)	1.598 (0.4)	-/-
G_α	9.335 (0.74)	9.45 (0.75)	-/-
P_τ	4.697 (0.5)	4.826 (0.47)	-/-
P_α	6.887 (0.51)	6.971 (0.5)	-/-

Notes: All variables are used as logarithms (i.e. are given in low letters). Westerlund (2007) presents four different panel cointegration tests with a null hypothesis of no cointegration. While G_τ and G_α test the alternative hypothesis of least one unit is cointegrated (i.e. group mean tests), P_τ and P_α test if the panel is cointegrated as a whole (i.e. panel mean tests). Short run dynamics are restricted to one lag and one lead. The "kernel with" is chosen according to the formula of 4(T/100)^2/9 and therefore to 3. All tests are implemented pairwise with a constant and a trend in the test regression. The robust p-values are given in parenthesis and are based on a bootstrapped distribution using 800 bootstrap replications in order to deal with cross-sectional dependencies.

Table C.7.3: Panel cointegration tests by Westerlund (2007); inequality vs. inequality
Annual data for 50 states plus DC from 1997-2011.

	top10	*top1*	*poverty*
top10			
G_τ		11.379 (0.02)	-0.657 (0.3)
G_α		11.830 (0)	7.23 (0.95)
P_τ		15.883 (0)	3.283 (0)
P_α		9.844 (0)	5.759 (0.93)
top1			
G_τ	-1.125 (0.01)		-3.718 (0.08)
G_α	10.012 (0.02)		8.102 (0.89)
P_τ	5.391 (0)		1.717 (0.02)
P_α	7.871 (0)		5.853 (0.76)
poverty			
G_τ	-0.939 (0.17)	7.475 (0.62)	
G_α	8.385 (0.11)	8.841 (0)	
P_τ	-0.956 (0.04)	7.059 (0.01)	
P_α	5.308 (0.04)	6.01 (0.01)	

Notes: All variables are used as logarithms (i.e. are given in low letters). Westerlund (2007) presents four different panel cointegration tests with a null hypothesis of no cointegration. While G_τ and G_α test the alternative hypothesis of least one unit is cointegrated (i.e. group mean tests), P_τ and P_α test if the panel is cointegrated as a whole (i.e. panel mean tests). Short run dynamics are restricted to one lag and one lead. The "kernel with" is chosen according to the formula of 4(T/100)^2/9 and therefore to 3. All tests are implemented pairwise with a constant and a trend in the test regression. The robust p-values are given in parenthesis and are based on a bootstrapped distribution using 800 bootstrap replications in order to deal with cross-sectional dependencies.

Table C.7.4: Dynamic OLS by Stock and Watson (1993) with lag(1) and lead (1) of first differences; economy vs. economy

Annual data for 50 states plus DC from 1997-2011.

	(w+l)-y	*y-l*	*y*	*w*	*l-p*	*w+l-p*
(w+l)-y						
L(0)	-/-	-0.442 *(-8.63)****	5.17 *(18.09)****	0.558 *(10.9)****	0.377 *(8.5)****	0.935 *(11.36)****
L(1)	-/-	-0.447 *(-8.35)****	5.16 *(17.37)****	0.553 *(10.34)****	0.354 *(7.66)****	0.907 *(10.54)****
y-l						
L(0)	-0.224 *(-8.36)****	-/-	0.682 *(2.71)****	0.776 *(28.95)****	0.294 *(9.2)****	1.07 *(21.20)****
L(1)	-0.216 *(-7.72)****	-/-	0.708 *(2.69)****	0.784 *(28.08)****	0.305 *(9.23)****	1.09 *(20.87)****
y						
L(0)	0.06 *(17.27)****	0.021 *(3.59)****	-/-	0.081 *(14.98)****	*0.01* *(1.91)**	*0.09* *(9.58)****
L(1)	0.06 *(16.52)****	0.022 *(3.56)****	-/-	0.082 *(14.49)****	*0.011* *(1.99)***	*0.092* *9.39****
w						
L(0)	0.275 *(10.97)****	*0.725* *(28.88)****	3.168 *(14.93)****	-/-	0.461 *(16.67)****	1.461 *(52.85)****
L(1)	0.28 *(10.53)****	*0.72* *(27.12)****	3.188 *(14.52)****	-/-	0.452 *(15.88)****	1.452 *(51.04)****
l-p						
L(0)	0.271 *(8.79)****	*0.395* *(9.27)****	0.612 *(2.11)***	0.667 *(17.06)****	-/-	1.667 *(42.66)****
L(1)	0.287 *(8.75)****	*0.393* *(8.66)****	0.711 *(2.35)***	0.68 *(16.68)****	-/-	1.68 *(41.19)****
w+l-p						
L(0)	*0.176* *(11.59)****	0.382 *(21.32)****	*1.363* *(9.68)****	*0.558* *(53.2)****	*0.442* *(42.07)****	-/-
L(1)	*0.182* *(11.23)****	0.381 *(19.96)****	*1.406* *(9.65)****	*0.563* *(52.3)****	*0.437* *(40.57)****	-/-

Notes: All variables are used as logarithms (i.e. are given in low letters). The bias corrected *t*-statistics of the coefficients are reported in parenthesis. * (**) [***] denotes that the coefficient is significantly different from zero at a 10% (5%) [1%] level. All equations include unreported, state-specific constants and a trend.

Table C.7.5: Dynamic OLS by Stock and Watson (1993) with lag(1) and lead (1) of first differences; ratio vs. ratio

Annual data for 50 states plus DC from 1997-2011.

	$(w_L+l_L)-(w+l)$	l_L-l	w_L-w
$(w_L+l_L)-(w+l)$			
L(0)	-/-	1.135 *(83.45)******	-0.135 *(-9.94)******
L(1)	-/-	1.124 *(80.91)******	-0.124 *(-8.95)******
l_L-l			
L(0)	0.809 *(81.4)******	-/-	-0.191 *(-19.17)******
L(1)	0.812 *(80.39)******	-/-	-0.188 *(-18.63)******
w_L-w			
L(0)	-0.965 *(-9.63)******	-1.965 *(-19.61)******	-/-
L(1)	-0.97 *(-9.47)******	-1.97 *(-19.24)******	-/-

Notes: All variables are used as logarithms (i.e. are given in low letters). The bias corrected *t*-statistics of the coefficients are reported in parenthesis. * (**) [***] denotes that the coefficient is significantly different from zero at a 10% (5%) [1%] level. All equations include unreported, state-specific constants and a trend.

Table C.7.6: Dynamic OLS by Stock and Watson (1993) with lag(1) and lead (1) of first differences; inequality vs. inequality

Annual data for 50 states plus DC from 1997-2011.

	top10	*top1*	*poverty*
top10			
L(0)	-/-	1.608 *(29.74)******	0.211 *(1.64)**
L(1)	-/-	1.459 *(27.55)******	0.15 *(1.12)*
top1			
L(0)	0.396 *(31.64)******	-/-	0.01 *(0.15)*
L(1)	0.43 *(33.35)******	-/-	0 *(0)*
poverty			
L(0)	*0.021* *(1.68)**	-0.001 *(-0.02)*	-/-
L(1)	*0.031* *(2.36)***	0.032 *(1.11)*	-/-

Notes: All variables are used as logarithms (i.e. are given in low letters). The bias corrected *t*-statistics of the coefficients are reported in parenthesis. * (**) [***] denotes that the coefficient is significantly different from zero at a 10% (5%) [1%] level. All equations include unreported, state-specific constants and a trend.

Table C.7.7: Weak Exogeneity in Cointegrated Panels by Moral-Benito and Servin (2014); economy vs. economy

Annual data for 50 states plus DC from 1997-2011.

	(w+l)-y	*y-l*	*y*	*w*	*l-p*	*w+l-p*
(w+l)-y	-/-	-2.109 *(1)*	-2.121 *(1)*	-2.125 *(1)*	-2.105 *(1)*	-2.113 *(1)*
y-l	-0.14 *(0.68)*	-/-	-0.12 *(0.68)*	-0.1 *(0.67)*	-0.138 *(0.68)*	-0.123 *(0.68)*
y	2.679 *(0.07)*	2.589 *(0.07)*	-/-	2.672 *(0.07)*	2.686 *(0.07)*	2.681 *(0.07)*
w	-0.92 *(0.92)*	-1.002 *(0.93)*	-1.073 *(0.95)*	-/-	-0.986 *(0.93)*	-1.037 *(0.94)*
l-p	-1.937 *(1)*	-1.933 *(1)*	-1.955 *(1)*	-1.939 *(1)*	-/-	-1.99 *(1)*
w+l-p	1.063 *(0.29)*	1.085 *(0.29)*	0.94 *(0.33)*	1.015 *(0.3)*	1.012 *(0.31)*	-/-

Notes: All variables are used as logarithms (i.e. are given in low letters). The null hypothesis is weak exogeneity of the cointegrated panel (i.e. of all US-states) by Moral-Benito and Servin (2014). The *p*-values are given in parenthesis.

Table C.7.8: Weak Exogeneity in Cointegrated Panels by Moral-Benito and Servin (2014); ratio vs. ratio

Annual data for 50 states plus DC from 1997-2011.

	$w_L+l_L)-(w+l)$	l_L-l	w_L-w
$w_L+l_L)-(w+l)$	-/-	0.349 *(0.51)*	0.05 *(0.61)*
l_L-l	-0.561 *(0.83)*	-/-	-1.01 *(0.94)*
w_L-w	*-2.251* *(1)*	-2.431 *(1)*	-/-

Notes: All variables are used as logarithms (i.e. are given in low letters). The null hypothesis is weak exogeneity of the cointegrated panel (i.e. of all US-states) by Moral-Benito and Servin (2014). The *p*-values are given in parenthesis.

Table C.7.9: Weak Exogeneity in Cointegrated Panels by Moral-Benito and Servin (2014); inequality vs. inequality

Annual data for 50 states plus DC from 1997-2011.

	top10	*top1*	*poverty*
top10	-/-	-2.697 *(1)*	-2.652 *(1)*
top1	-2.611 *(1)*	-/-	-2.598 *(1)*
poverty	*-2.331* *(1)*	-2.232 *(1)*	-/-

Notes: All variables are used as logarithms (i.e. are given in low letters). The null hypothesis is weak exogeneity of the cointegrated panel (i.e. of all US-states) by Moral-Benito and Servin (2014). The *p*-values are given in parenthesis.