The Farm, the City, and the Emergence of Social Security*

Elizabeth M. Caucutt[†], Thomas F. Cooley[‡], and Nezih Guner[§]

November 23, 2012

Abstract

We study the social, demographic and economic origins of social security. The data for the U.S. and for a cross section of countries suggest that urbanization and industrialization are associated with the rise of social insurance. We describe an OLG model in which demographics, technology, and social security are linked together in a political economy equilibrium. In the model economy, there are two locations (sectors), the farm (agricultural) and the city (industrial) and the decision to migrate from rural to urban locations is endogenous and linked to productivity differences between the two locations and survival probabilities. Farmers rely on land inheritance for their old age and do not support a pay-as-you-go social security system. With structural change, people migrate to the city, the land loses its importance and support for social security arises. We show that a calibrated version of this economy, where social security taxes are determined by majority voting, is consistent with the historical transformation in the United States.

Keywords: Social Security, Political Economy, Structural Change, Migration

JEL Classifications: H55, H3, D72

*We thank Jeff Campbell, Dirk Krueger, Per Krusell, Lance Lochner, Stephane Pallage, and seminar and conference participants at Canadian Macro Study Group Meetings in Montreal, CEMFI, CEPR Annual Public Policy Symposium in Kiel, Federal Reserve Bank of Chicago, Federal Reserve Bank of Cleveland, Indiana University, REDg General Equilibrium Macroeconomics Workshop in Santiago de Compostela, Society for Economic Dynamics Meetings in Prague, Universitat Pompeu Fabra and University of Rochester for comments.

†Department of Economics, Social Science Centre, The University of Western Ontario, London, ON N6A 5C2, Canada.

‡Stern School of Business, New York University, 44 West Fourth St., Suite 11-58, New York, NY 10012-1126, USA and NBER.

§ICREA-MOVE, Universitat Autonoma de Barcelona, Barcelona GSE. MOVE Facultat d'Economia Edifici B – Campus de Bellaterra, Universitat Autonoma de Barcelona, 08193 Bellaterra, Cerdanyola del Vallès, Spain.

1 Introduction

The late 19th and early 20th centuries in the Unites States were characterized by a movement from a primarily rural and agricultural economy to a primarily urban and industrial economy. Figure 1 shows the geographic distribution of the U.S. population between 1800 and 1940.[1] In the beginning of the 19th century 94% of the population was living in rural areas. By 1940 the share of population living in rural areas was 43.5%, while the share living on the farm was only 23%. Coincident with this dramatic shift in the structure of the economy came changes in the institutional needs of the population. The sorts of social care arrangements that were common place on the farm were harder to implement and enforce in the city, and the shifting population gave rise to new political coalitions with disparate views on social policy. Many prominent accounts of changing institutions of this period, e.g. Wiebe (1966), Sass (1997) and Schieber and Shoven (1999), emphasize the critical role that urbanization and industrialization played in the creation of new institutions: "The willingness of the U.S. to finally go the route of so many other countries in adopting a national social insurance program in 1935 was the result of three major forces. The first was the increased dependence on wage income that had arisen over the preceding half-century as the country had industrialized," (Schieber and Shoven 1999, page 18). Indeed, the Social Security Administration (2003) characterizes the year 1920 as "a historical tipping-point. In that year, for the first time in our nation's history, more people were living in cities than on farms." Of course there were other forces besides industrialization: "The second force was the terrible economic environment caused by the Great Depression. And the third was a complicated set of political movements raising fundamental questions about economic and political structures we had adopted," (Schieber and Shoven 1999, page 18). Although the Great Depression is often considered a major impetus for the social security legislation in the U.S., its effects are far from clear. Miron and Weil (1998) conclude their study on the origins of social security by stating that: "Regarding the lasting impact of the Great Depression, our conclusion is that there was surprisingly little," (page 321).[2] Furthermore, although the Great Depression stimulated some support for social security as a means to improve the conditions of struggling poor (Temin 1991), the fact that many industrialized countries introduced a social security system before the Great Depression suggests that other, more fundamental forces must be in play. The contribution of this paper is to provide the first economic mechanism that can lead

[1] Appendix A provides data sources for all figures and tables.

[2] On the macroeconomic effects of Great Depression, see Cole and Ohanian (2004).

social security to emerge as a direct result of simultaneous urbanization and industrialization. While ours may not be the only mechanism at work, we demonstrate in a simple, empirically plausible framework that it is consistent with the experience of the United States prior to, and immediately after the introduction of social security.

We propose that the rural (agricultural) to urban (industrial) shift is a critical factor explaining the emergence of social insurance, more specifically, social security. The correlation between urbanization and social insurance has been recognized by political scientists and sociologists.[3] Figure 2 shows the level of urbanization and the fraction of the elderly (65+) population across U.S. states in 1930. About 23 states (those encircled) introduced a state pension plan before the 1935 Social Security Act. Of those, 18 states had an urbanization rate higher than the U.S. average. The correlation between the fraction of elderly population and state pensions is also positive but smaller than that for urbanization (only 65% of states with higher than average elderly population had adopted pension plans). Although this picture provides only suggestive evidence, the basic relation seems to hold up to closer empirical scrutiny. Amenta and Carruthers (1988) look at the timing of old age pension plan adoption among U.S. states. They find a statistically significant effect of urbanization on the passage of old age pension plans. More compelling is that the relationship remains in cross-country data. Figure 3 shows the correlation between the fraction of the labor force in agriculture at the start of the 20th century and the date in which a social security system was introduced among European countries. Clearly, a larger labor force in agriculture is associated with later adoption. Kim (2001) investigates the timing of old-age pension adoption across O.E.C.D. countries in more detail and finds that the percent of labor not employed in agriculture is strongly associated with the adoption of old-age pensions.

Figure 3 suggests that social security was adopted relatively late in the U.S. Furthermore, according to Figure 2, many of the states in the U.S. with low urbanization were Southern states where agricultural interests and land inequality was stronger. Galor, Moav and Vollrath (2009) show that land inequality as a reflection of strong agricultural interest was a deterrent to the adoption of human capital promoting institutions in the U.S., such as public schooling and child labor regulations. They demonstrate that this lack of investment reduced the mobility of workers and slowed the transition from an agricultural to an urban economy. Their framework highlights how the distribution of land is important for the pace of industrialization. In a similar vein, one reason for the late adoption of social security (which is more attractive for industrial

[3] See eg. Pryor (1968) and Collier and Messick (1975)

workers) might have been the opposition by Southern states in the U.S. Wealthy land holders, who had little need for social security, benefited from the lack of such programs as they reduced the incentive for their workers to migrate. Indeed, the struggle over the introduction of social security might also reflect the divergent interests of agricultural and industrial elites, which we abstract from in the current analysis.

We argue that industrialization combined with demographic shifts can account for the dramatic change in the social insurance system in the United States. Farmers rely on land inheritance for their old age. As a result, they do not support a pay-as-you-go social security system since they have little incentive to save to start with. With structural change, people migrate to the city and the land loses its economic value. Social security becomes an attractive way to shift resources to older ages and support for social security arises. To formalize this argument, we develop an overlapping generations economy with two sectors, which we interpret as agricultural and industrial. Farm production requires capital, labor and land. Land is a fixed factor, so there are decreasing returns to labor. City production on the other hand requires capital and labor and exhibits constant returns to scale.[4] Agents in this economy live up to three periods, as young, middle aged and old. They face an exogenous probability of dying at the end of the second period of their lives. In each period agents earn wages in the sector in which they are located. We assume that they can save, but are unable to borrow. Land is passed from one generation to another by inheritance, as are accidental capital bequests. Each period young agents make a once and for all decision about where to live.[5] There is also a social security mechanism that can tax the young and the middle aged and pay transfers to the old.

The savings and optimal tax decisions of agents in this environment imply that an agent will support social security if the following two conditions hold: First, the return to social security

[4] Hansen and Prescott (2002) model the industrial revolution as a switch from a (Malthus) production technology with a fixed factor of production, land, to a (Solow) production technology, with no fixed factors. Parente and Prescott (2005) use a similar framework to study the evolution of international income levels since 1750. Galor and Weil (2000) provides a framework in which transition from stagnation to growth occurs endogenously within a unified framework. Laitner (2000) studies a two-good, two-sector model in which, like Hansen and Prescott (2002), only the agricultural sector uses land. He investigates why the savings rate increases with development. Other well-known models of structural change are Echevarria (1997) and Kongsamut, Rebelo and Xie (2001). Greenwood and Seshadri (2002) and Gollin, Parente and Rogerson (2002) model the shift of labor from agriculture to manufacturing, and the associated pattern of rural to urban migration that is associated with process of economic development.

[5] Among recent models with an explicit location decision see Vandenbroucke (2008), Hassler, Rodríguez Mora, Storesletten and Zilibotti (2005), and Klein and Ventura (2009).

must be greater than the return to private assets. For a middle-aged agent, the return to social security is relatively high because he views the cost as one period of tax payments, while his benefits next period are based on two generations worth of payments. There is also an annuity component to social security, where the returns are divided amongst the surviving agents. Second, the agent must want to save for old age. Whether or not the middle-aged individual wants to save for old age depends crucially on his age-income profile. If his income profile is steep, i.e. he expects to earn substantially more in old-age, he is unlikely to want to sacrifice resources in middle age for even more consumption in old age. In our framework, those middle-aged farmers who have a surviving parent do not inherit land while middle aged. However, they inherit land in old age, conditional on survival. Both middle-aged farmers who own land and middle-aged-city workers have relatively flatter age-income profiles, and are more likely to support social security. We show that in a situation where everyone is living on the farm, if the age-income profile of middle-aged-landless farmers is steep enough, a majority of the population will oppose the implementation of social security. These individuals prefer to rely on their land inheritance for old-age security. Once enough rural-urban migration takes place, a majority of the population will no longer be in a position to inherit land, and social security is implemented.

In order to quantify whether our story is consistent with the experience of the United States before (1800) and after (1940) social security is introduced, we need to take a stand on how preferences over tax levels are aggregated in a political process. In doing so, we merge two literatures: the political economy of social security and the economics of structural change. This allows us to study the set of demographic, social, and economic conditions that give rise to an economy without social security and the changes that would eventually lead to the introduction of publicly managed old age security. We assume that the level of the social security tax is determined by majority voting.[6] In the initial steady state of this economy the relative productivity of the farm sector is high and survival probabilities are low. As a result, farm incomes are high relative to city incomes. All agents live on the farm, and land is an important source of income for the old. The median voter is a middle-aged farmer who does not own land and who prefers a zero social security tax. Once the city is more productive and individuals live longer, people migrate, and the importance of land diminishes. In the final steady state, the median voter is a middle-aged-city worker who prefers a positive social security tax. While the framework is relatively simple,

[6] The current paper follows the recent literature on dynamic models of political economy; see among others Krusell, Quadrini, and Ríos-Rull (1997), Krusell and Ríos-Rull (1999), Hassler, Rodríguez Mora, Storesletten and Zilibotti (2003), and Corbea, D'Erasmo and Kuruscu (2009).

it leads to a rich political economy environment. The identity of the median voter is not just age, but also location, which turns out to be critical for generating the emergence of social security. This is achieved by merging the structural transformation from farm to city with the political economy of institutions, in this case social security. Hence, the current paper is closely related to the literature that links structural changes with changes in political and social institutions.[7]

There is a large literature on the political economy of social security systems that analyzes the political sustainability of Pay-as-you-go (PAYG) social security.[8] The conclusion of most of this literature is that support for social security in democratic societies depends on the age of the median voter.[9] These papers are oriented toward explaining why an existing system can survive, expand or shrink.[10] They cannot address why such a system was started in the first place, or more precisely, why such systems have not always existed.[11] By allowing the identity of the median voter to include his geographical location, we overcome this shortcoming and provide a framework in which the emergence of social security is a response to the urbanization of the

[7] Doepke and Zilibotti (2005) study how skilled bias technological progress, which lowers fertility and increases the importance of education, can lead to the adoption of child labor laws. Galor and Moav (2006) show that it might be in capitalists' own interest to expand public education to the masses as a result of the growing importance of human capital for the production process. Doepke and Tertilt (2009) show that the growing importance of human capital can also trigger men to grant political rights to women. Galor et al (2009) study the effects of the concentration of land ownership on human capital accumulation and growth within a political economy model. Bertocchi (2011) studies the long run decline in the importance of bequest taxes within a two-sector (agriculture and manufacturing) dynamic political economy model. In her model, land is easier to tax than capital. The decline of agriculture, which reduces the value of land, makes bequest taxes an unattractive option over time.

[8] There also exists a large literature that analyzes macroeconomic and distributional implications of the current social security system without political economy considerations (e.g. Imrohoroglu, Imrohoroglu and Joines 1985).

[9] Cooley and Soares (1999), Galasso (1999), and Boldrin and Rustichini (2000) build models in which non-altruistic median voters decide to keep an existing system. The median voter's decision depends on two factors in these models: First, there exists a reputational mechanism in place which eliminates all future benefits if the median voter deviates from the current arrangement. Therefore, a median voter cannot avoid taxes today and hope to get benefits in the future. Second, the median voter might want to keep an existing social security system in order to benefit from the high interest rates associated with a depressed capital stock.

[10] For example, Cooley and Soares (1996) study an economy in which the initial generation votes over a social security replacement rule that depends on the age structure of the population. Hence, as the population structure changes (e.g. as a result of the Baby Boom) a rule that was sustainable in the past can become unsustainable. Gonzalez-Eiras and Niepelt (2005) link the size of intergenerational transfers to the age structure of the population. Conesa and Krueger (1999) study how the status-quo bias is related to idiosyncratic uncertainty.

[11] Krueger and Kubler (2005) study how the introduction of an unfunded social security system can be Pareto improving in an economy with incomplete markets.

population.

1.1 Facts

What were the economic and demographic forces that led to this shift from rural to urban population? One obvious answer to this question is the increase in the city wage relative to the farm wage that arose from greater technical change in the city relative to the farm. GDP per person employed increased by a factor of 3.5 in the U.S. between 1870 and 1940 (Maddison 2001). While productivity in both the agriculture and the non-agricultural sectors grew rapidly during this period, the growth in non-agricultural sectors was faster than the growth in agriculture, leading to the transformation of the U.S. economy (see Greenwood and Seshadri 2002 and Greenwood and Uysal 2005). Figure 4 shows the change in total factor productivity (TFP) in agricultural and non-agricultural sectors in the United States. Between 1800 and 1940, TFP grew by a factor of 1.92 in agriculture, while it grew by a factor of 4.21 in manufacturing.

Another possible impetus for rural-urban migration is the increase in life expectancy that took place over this time period. As life expectancy increased, two important changes occurred in the agricultural sector. First, the amount of farm labor relative to farm land rose, causing farm wages to fall. Second, as farmers lived longer, the transfer of land ownership via inheritance was delayed. Both events increased the relative attractiveness of living in the city for farmers, and encouraged rural-urban migration. Of crucial importance for this story is not that life expectancy at birth increased, but that life expectancy conditional on reaching or getting near retirement age increased. Figure 5 shows the changes in conditional survival probabilities from age 60 to 65, from 65 to 70, from 70 to 75, and from 75 to 80. Survival probabilities increased by about 5 percentage points between 1850 and 1900, and by another 2 percentage points between 1900 and 1940.

What were the key features of the 19th century farm economy? First, the old in the 19th century had relatively more wealth than the old in the 20th century and land as an illiquid asset provided an important source of income and wealth for the elderly. In 1850, those 60 years or older had about three times as much real estate wealth as the 30-39 age group (see Williamson and Lindert 1980, Table 1.7) and an analogous picture emerges for total wealth in 1870 (see Soltow 1992, Table 3.2). It is also true that today agents around retirement age, 65-69, have high levels of wealth. Their relative position, however, has deteriorated. Typical age-wealth profiles from the 19th century show a continuous rise until age 70 and a slight decline afterwards. By 1962, the age-wealth profiles had two peaks, one for ages 55-59 and another for 65-69. By 1982, the first peak occurred for even younger ages, 45-49 (Wolf 1992). It is therefore not surprising that Schieber and Shoven (1999) conclude that the over-65 age cohort controlled more wealth than any other group in the early 19th century.

Second, the land was illiquid. Land was very slow to sell, people had to use costly advertisements in the newspaper or word of mouth. Land, unlike other securities, was useless unless it was worked or rented, so it wasn't a preferred collateral. Mortgaging land was difficult; "it almost never seemed clear who owned what land, how much that land was worth, and whether prospective borrowers had mortgaged the land elsewhere" (Wright 2001, page 28).

Third, inheritance, and in particular inheritance of land, played a key role in wealth accumulation. According to Soltow (1982) inheritance was the determining factor of wealth inequality in the U.S. during the 19th century. Inheritance was a much more significant factor than life-cycle savings in explaining the relationship between age and wealth in the U.S. in 1870. Overtime, with longer life expectancy and sustained economic growth, the importance of inheritance declined. DeLong (2003) estimates that in pre-industrial times inheritance contributed around 90% to wealth acquisition of a cohort, while the share of inheritance is less than 50% today. The situation in the 19th century America, with its dynamic economy and emphasis on equal division of bequests, was likely less dramatic than pre-industrial times, but still the role of inheritance must have been much bigger than today. Moreover, land was the most important form of inheritance in the 19th century.[12] The farm population consisted mainly of workers and owner-farmers. Renting the land to others was not common. According to Yang (1992), about 90% of farmers were owners in 1860.[13] Young and middle-age workers without land look forward to inheriting land when old and therefore see little need for other savings vehicles.[14]

Finally, the long-term borrowing opportunities were very limited, if not non-existent. Informal credit networks in which farmers and artisans borrowed from merchants and from each other were active even before the establishment of banks and other formal credit institutions (Rothenberg 1992). These informal credit arrangements, however, mainly served short term credit needs of

[12] In his study of Butler County (Ohio), Newell (1986) documents that for the 1803-1865 period, inheritance consisted almost exclusively of real property.

[13] Even at the end of the 19th Century, most farmers were owners, see Barlowe and Timmons (1950).

[14] As Sass (1997, page 5) points out "The family enterprise institution also vested the old with powerful property rights vis-a-vis their adult children. Elderly parents held first claim on the firm and its assets, while their offspring remained dependent for their incomes and inheritance... parents retained ownership over the main body of family assets and chose when they would transfer farms and businesses to their children."

farmers, and did not provide them with enough credit to smooth their life-cycle consumption. Hence, while middle-aged farmers could borrow small amounts for a few months, they were a long way from borrowing against the land they might acquire in the future. Establishment of formal banks did not change this picture in a fundamental way.

The Bank of North America, the first bank in the modern sense, was established in 1781 in Philadelphia. The following decades witnessed an increase in the number of Banks from about 28 in 1800 to 824 in 1850 (Carter et al 2006, Table Cj142-148). Despite this increase in the number of banks, the financial sector did not serve the long-term credit needs of agriculture. The Bank of North America was a strictly commercial bank. The tension between the short term commercial needs of merchants and the long term credit needs of farmers was a key aspect of the early U.S. banking. The early banks, with merchants as their main stockholders, were mainly engaged in providing short term commercial credit. As Table I documents, the maturity of credit was very short and the amount of credit was small. Per capita loans and discounts by banks was about 16% of GDP per capita in 1850 and the maximum maturity was 6 months. Not surprisingly, a large segment of the population did not directly deal with a bank.[15]

Table I - Banks and Credits

Year	Population (thousand)	GDP per capita	Loans and Discounts by Banks (thousand $)	Loans and Discounts (per capita)	Credit Maturities
1800	5297	98			30-45 days
1820	9618	86			60-90 days
1834	14504	86	324119	22	6 months
1840	17120	101	462897	27	6 months
1850	23261	109	364204	16	6 months

Wang (2008) studies the distribution of borrowers from Plymouth Bank in Plymouth County, Massachusetts. This was the first bank in Massachusetts, and its practices show that the introduction of banks did not broaden access to credit. In the early 1800s, about 60% of bank customers for credit were merchants, while farmers represented only 8%. During that same period,

[15] This was true even for financially developed cities like New York. "Most early New Yorkers did not have a bank account. In 1825 New York contained about 60,000 people per solvent commercial bank. By 1835 this number had dropped to about 26,000. But even in 1855, when New York had 11,000 people per bank, ... it is clear fairly large segments of the population did not have direct dealings with New York Banks" (Wright 2001, page 114).

more than half (54%) of the adult population in Plymouth county were engaged in agriculture, in contrast to merchants who were only 12%. By 1850, there was not much improvements in the representation of farmers among credit customers.[16]

Even when mortgages were available in later years, "only available mortgages were short-term, balloon mortgages. Such loans are unamortized. Periodic payments meet the interest but contribute nothing to the principal, which is payable in full at maturity. Mortgages typically lasted three years or less and might be renewed, though renewal conditions were never certain...As late as 1890, only 29% of farmers were encumbered by mortgages, and among those that were, the debt average about 35% of their worth." (Atack et al 2000, page 275).

These conditions led to the populist movement of the late 19th century, which responded to the fact that farmers were credit constrained, and placed the blame on Wall Street and Eastern bankers. In the early 20th century, the government began various agricultural credit programs (such as the Federal Farm Loan Act of 1916) to try to alleviate these problems but they were of limited success. The problems of inadequate access to credit for farmers continued into the early 20th century.

In the next section we describe the economic environment and the recursive competitive equilibrium given an exogenous political process. We characterize agents' decisions, given fixed prices in Section 3. Here, we also demonstrate analytically how an economy shifts from an agrarian population with no social security system to a primarily urban population with a social security system. In Section 4 we describe how taxes are determined. We discuss the results of our quantitative exercises in Section 5, and conclude in Section 6.

2 Environment

Consider the following one-good, two-sector overlapping generations model. In the first sector (or location), which we will call the *farm*, capital, labor and land are combined to produce output. In the second sector (or location), which we call the *city*, the same good is produced using capital and labor.

Agents live a maximum of 3 periods, which we refer to as young, middle-aged and old, and face a probability, π, of surviving from the second to the last period. Let β be the discount

[16] According to Wang (2008, page 446): "Namely, farmers and artisans did not have easy access to banks. They usually borrowed in the personal credit market." (page 446). "Thus despite a well-developed market, most potential borrowers in Massachusetts did not have access to bank credit." (page 456).

factor. The objective of a young person is to maximize

$$U(c_y, c_m, c_o) = u(c_y) + \beta u(c_m) + \beta^2 \pi u(c_o), \tag{1}$$

where c_i, $i \in \{y, m, o\}$, denotes age-i consumption, and u is continuous, strictly increasing and strictly concave.

Each period every middle-aged person has a child who is born into their parent's location. When an agent is born on the farm, he makes a once-and-for-all decision to stay there or move to the city. Those who are born in the city are not allowed to move to the farm. Middle-aged and old agents can't change their location.[17] Let the fraction of young, middle-aged and old agents who live on the farm be denoted by λ_y, λ_m and λ_o, respectively.

In both locations young, middle-aged and old all inelastically supply one unit of labor.[18] Agents are endowed with location dependent efficiency units ε_i^j, $j \in \{f, c\}$ and $i \in \{y, m, o\}$. Since only a fraction π of middle-aged people survive to old age, the total labor supply on the farm is given by $N^f = \lambda_y \varepsilon_y^f + \lambda_m \varepsilon_m^f + \lambda_o \varepsilon_o^f \pi$ and the total labor supply in the city by $N^c = (1-\lambda_y)\varepsilon_y^c + (1-\lambda_m)\varepsilon_m^c + (1-\lambda_o)\varepsilon_o^c \pi$. Agents are located either in the city or on the farm and can only work in that sector. There is a competitive labor market in each location. Let w^j denote the wages in sector j. The labor income of an age-i agent in location j is $w^j \varepsilon_i^j$ for $i \in \{y, m, o\}$ and $j \in \{f, c\}$.

People are not allowed to borrow, but can accumulate capital and rent it to firms in either sector at a competitive rate, ρ. Capital moves costlessly between the farm and the city, so let $r = \rho - \delta$ be the common rate of return to capital, where $\delta \in [0, 1]$ is the common rate of capital depreciation. There is no market in which agents can buy and sell land. Each agent is born without any assets (including land). On the farm, when an agent dies (at the end of the second or third period), his land is inherited by the oldest surviving descendant. Therefore, a fraction of the land is owned by the $\pi\lambda_o$ surviving old, and the remainder is owned by the $(1-\pi)\lambda_m$ middle-aged who inherited land early. Below we refer to middle-aged farmers with land as landed farmers and those without land as landless farmers. We normalize the total amount of land to 1, so each landholding farmer has $\frac{1}{\pi\lambda_o + (1-\pi)\lambda_m}$ units of land. Farmers rent their land to firms at a competitive rate q.

[17] The vast majority of migration from the farm to the city consisted of young workers. (Schieber and Shoven (1999), p. 18, and U.S. Bureau of the Census (1975), pp. 139, 465)

[18] We therefore abstract from the rise in retirement (i.e. decline in the labor force participation of old) since the 1850s. See Kopecky (2011) for a model with endogenous retirement that links this rise to the technological progress in the production of leisure goods.

In both locations some middle-aged agents receive accidental capital bequests from their parents. As a result, middle-aged agents differ in their asset and land holdings on the farm, while they only differ by their asset levels in the city. If a young farmer chooses to move to the city, he gives up all claims on his parent's land, and that land, upon his parent's death, is divided equally among the remaining land owners (we show later that relaxing this assumption does not change the main results). However, he still receives any accidental asset bequest his parent might leave, as we assume capital can freely move between the farm and the city.

Each sector is populated by a large number of production units (family farms in the agricultural sector and factories in the city sector) which have access to constant returns to scale production functions represented by

$$Y^f = \gamma^f F^f(K^f, N^f, L), \tag{2}$$

and

$$Y^c = \gamma^c F^c(K^c, N^c), \tag{3}$$

where variables Y^j, K^j, N^j and L, $j \in \{f, c\}$, refer to output, capital, and labor employed in each sector, and land used in the farm sector, respectively. The parameter γ^j, $j \in \{f, c\}$, is the total factor productivity (TFP) in sector j. Land is a fixed factor and used only in the farm sector.

Given the wage rate in sector j, w^j, the rental rate for capital, ρ and the rental rate for land, q, the problem of a production unit in the farm sector is given by

$$\max_{N^f, K^f, L} \left\{ Y^f - w^f N^f - \rho K^f - qL \right\},$$

subject to (2), and in the city sector by

$$\max_{N^c, K^c} \left\{ Y^c - w^c N^c - \rho K^c \right\},$$

subject to (3).

Finally, there is an economy-wide social security system that collects a lump-sum tax, τ, from the young and the middle-aged and provides each old with an amount $2\tau/\pi$.

2.1 Recursive Equilibrium

At any point in time, the aggregate state in this economy consists of the distribution of capital across agents, the distribution of agents across the city and the farm, and the prevailing tax level. Since agents are born without any capital, all capital is owned by the middle-aged and the old. Because they make different decisions, we differentiate between the asset distribution of those in the city and those on the farm, and between the farmers with land and the farmers without land. We represent the distribution of capital across old city and farm residents by ψ_o^c and ψ_o^f, and middle-aged city residents and farmers by ψ_m^c and $\psi_m^{f\kappa}$, with $\kappa = 1, 0$ indicating whether a middle-aged farmer is landed, $\kappa = 1$, or landless, $\kappa = 0$. In what follows we let $\Psi = (\psi_m^c, \psi_o^c, \psi_m^{f1}, \psi_m^{f0}, \psi_o^f)$ be the set of asset distributions. We represent the distribution of agents between the two locations, city and farm, by $\Lambda = (\lambda_y, \lambda_m, \lambda_o)$, where λ_j is the fraction of age-j agents who live on the farm.

Let, $S = (\Psi, \Lambda, \tau)$ be the aggregate state. The evolution of the aggregate state is given by two functions, G and H, where $\Psi' = G(S)$ is next period's asset distribution and $\Lambda' = H(S)$ is next period's distribution of agents across locations, given the current state S. When individuals solve their problems, they take the transition functions G and H as given. However, G and H, must be consistent with individual decisions in equilibrium. In online Appendix B, we analyze how the savings and location decisions of agents determine the evolution of the asset distribution and the fraction of agents living in each location. Below, we first describe the recursive competitive equilibrium given a constant tax level, τ. In Section 4, we allow τ to be determined and possibly change through majority voting in a political economy equilibrium.

2.1.1 City Problem

We begin by describing the economic problem of agents in the city. We approach agents' problems recursively, starting from the problem of an old agent, whose state consists of the aggregate state, $S = (\Psi, \Lambda, \tau)$, and his individual asset level a. Let $V_o^c(a, S)$ denote the value of being an old person with asset level, a. Since the old will simply consume their total resources, this is given by

$$V_o^c(a, S) \quad = \quad u(w^c \varepsilon_o^c + (1 + r)a + \frac{2\tau}{\pi}), \tag{4}$$

where for expositional clarity we suppress the dependence of r and w^c on the aggregate state S.

Next, we look at the decisions of middle-aged agents. Their decisions are determined by

$$
\begin{aligned}
V_m^c(a,S) &= \max_{a'} \left\{ u(w^c \varepsilon_m^c + (1+r)a - \tau - a') + \beta\pi V_o^c(a',S') \right\}, \quad (5)\\
s.t.\ S' &= (G(S), H(S), \tau),\\
a' &\geq 0,
\end{aligned}
$$

where next period's distribution of assets across agents, Ψ', is given by $G(S)$, and the distribution of agents between the two locations, Λ', is given by $H(S)$. Let $a_m^c(a,S)$ denote the savings decision of a middle-aged-city person with individual asset level, a, that results from problem (5).

Finally, we consider the decisions of the young agents who are born in the city. Let $b(a,S)$ denote the bequest a young agent receives if his middle-aged parent starts middle-age with assets, a, and dies before reaching old age. The problem of a young agent is then given by

$$
\begin{aligned}
V_y^c(b(a,S),S) &= \max_{a'} \{ u(w^c \varepsilon_y^c - \tau - a') + \beta\pi V_m^c(a',S') \quad (6)\\
&\quad + \beta(1-\pi)V_m^c(a' + b(a,S), S')\},\\
s.t.\ S' &= (G(S), H(S), \tau),\\
a' &\geq 0.
\end{aligned}
$$

Let $a_y^c(b(a,S),S)$ be the savings decision of a young agent who expects to get $b(a,S)$ as a bequest next period.

2.1.2 Farm Problem

The value function for an old agent on the farm is similar to the old agent's in the city, except the old farmer earns land income. His value function is given by

$$
V_o^f(a,S) = u\left(w^f \varepsilon_o^f + (1+r)a + \frac{q}{\pi\lambda_o + (1-\pi)\lambda_m} + \frac{2\tau}{\pi} \right), \quad (7)
$$

where $\frac{1}{\pi\lambda_o+(1-\pi)\lambda_m}$ is the per capita amount of land on the farm, and as in equation (4), we suppress the dependence of prices, w^f, r, and q, on S.

The problem of middle-aged agents on the farm differs from that of those in the city, because middle-aged farmers differ in land-holding status. Let $\kappa = 0$ if the farmer is landless, and let

$\kappa = 1$ if the farmer is landed. The middle-aged farmer's problem can be written, for $\kappa = 0, 1$, as

$$\begin{aligned} V_m^{f\kappa}(a,S) &= \max_{a'}\{u\left(w^f \varepsilon_m^f + (1+r)a + \frac{q\kappa}{\pi\lambda_o + (1-\pi)\lambda_m} - \tau - a'\right) \\ &\quad + \beta\pi V_o^f(a', S')\}, \\ s.t.\ S' &= (G(S), H(S), \tau). \\ a' &\geq 0. \end{aligned} \tag{8}$$

Let $a_m^{f\kappa}(a,S)$ be the decision rule for middle-aged farmers. Note that a middle-aged-landed farmer who survives to old age may have a different level of land holdings than he has today. Land per farmer may change due to migration, since migration alters the distribution of agents across locations, which is captured by $\Lambda' = H(S)$.

When considering the young farmer's savings decision, we need to do so jointly with his location decision. His saving decision will depend on where he chooses to live. First, consider a young farmer who stays on the farm. If his parent dies next period, he will receive an accidental capital bequest. The amount will depend on his parent's savings decision, which depends on the land holding status of the parent. Therefore, although the young do not differ by asset level or land holding, we label them with their parent's asset and land holding status. In particular, let $b^\kappa(a,S)$ denote the capital bequest that a young agent gets upon the early death of his parent, who has a units of capital and land holding status $\kappa = 0, 1$. Note that if his parent dies, he will also receive the land his parent leaves behind. A young agent who decides to stay solves

$$\begin{aligned} V_y^{f\kappa s}(b^\kappa(a,S), S) &= \max_{a'}\{u(w^f \varepsilon_y^f - \tau - a') + \beta\pi V_m^{f0}(a', S') + \\ &\quad \beta(1-\pi)V_m^{f1}(a' + b^\kappa(a,S), S')\}, \\ s.t.\ S' &= (G(S), H(S), \tau), \\ a' &\geq 0. \end{aligned} \tag{9}$$

Let his savings decision be represented by $a' = a_y^{f\kappa s}(b^\kappa(a,S), S)$.

Next consider a young farmer who *goes* to the city. If his parent dies, he will only receive a capital bequest of $b^\kappa(a,S)$. He solves

$$
\begin{aligned}
V_y^{f\kappa g}(b^\kappa(a,S),S) &= \max_{a'}\{u(w^c\varepsilon_y^c - \tau - a') + \beta\pi V_m^c(a',S') + \\
&\qquad \beta(1-\pi)V_m^c(a'+b^\kappa(a,S),S')\}, \\
s.t.\ S' &= (G(S),H(S),\tau), \\
a' &\geq 0.
\end{aligned} \tag{10}
$$

Let his savings decision be given by $a' = a_y^{f\kappa g}(b^\kappa(a,S),S)$.

Finally, let $L(b^\kappa(a,S),S)$ be an indicator of whether the young farmer is a goer or a stayer, which is simply determined by comparing his expected lifetime utility in each location, i.e.

$$
L(b^\kappa(a,S),S) = \begin{cases} 1, \text{ if } V_y^{f\kappa g}(b^\kappa(a,S),S) \geq V_y^{f\kappa s}(b^\kappa(a,S),S) \\ 0, \text{ otherwise} \end{cases}. \tag{11}
$$

2.1.3 Equilibrium

Given a policy τ, a *recursive competitive equilibrium* for this economy consists of a set of value functions, $V_y^c(b(a,S),S)$, $V_m^c(a,S)$, and $V_o^c(a,S)$, for agents who live in the city and $V_y^{f\kappa s}(a,S)$, $V_y^{f\kappa g}(a,S)$, $V_m^{f\kappa}(b^\kappa(a,S),S)$, $\kappa = 0,1$, and $V_o^f(a,S)$ for agents who live on the farm; a set of decision rules $a_y^c(b^c(a,S),S)$ and $a_m^c(a,S)$ for agents who live in the city, and $a_y^{f\kappa s}(b^\kappa(a,S),S)$, $a_y^{f\kappa g}(b^\kappa(a,S),S)$ and $a_m^{f\kappa}(a,S)$, $\kappa = 0,1$, for agents who live on the farm; a location rule for young farmers, $L(b^\kappa(a,S),S)$, $\kappa = 0,1$; a set of pricing functions $r(S)$, $w^c(S)$, $w^f(S)$, and $q(S)$, and aggregate laws of motion $G(S)$ and $H(S)$ such that:

- Given the transition functions $G(S)$ and $H(S)$, and pricing functions $r(S)$, $w^c(S), w^f(S)$, and $q(S)$, the value functions and corresponding decision rules solve the appropriate household problems in equations (4), (5), (6), (7), (8), (9), (10), and (11), with $b(a,S) = a_m^c(a,S)$ and $b^\kappa(a,S) = a_m^{f\kappa}(a,S)$, $\kappa = 0,1$.

- The pricing functions, $r(S)$, $w^c(S), w^f(S)$, and $q(S)$, are determined by profit maximization of production units in each sector together with a no arbitrage condition for capital, i.e. $r(S)$, $w^c(S), w^f(S)$, and $q(S)$ satisfy

$$
w^c(S) = F_2^c(K^c, N^c),
$$

$$
w^f(S) = F_2^f(K^f, N^f, L),
$$

$$q(S) = F_3^f(K^f, N^f, L),$$

and,

$$r(S) + \delta = F_1^c(K^c, N^c) = F_1^f(K^f, N^f, L),$$

with aggregate labor and capital in each sector given by

$$N^f = \lambda_y \varepsilon_y^f + \lambda_m \varepsilon_m^f + \lambda_o \pi \varepsilon_o^f,$$

$$N^c = (1 - \lambda_y)\varepsilon_y^c + (1 - \lambda_m)\varepsilon_m^c + (1 - \lambda_o)\pi \varepsilon_o^c.$$

and,

$$K = K^c + K^f = A^c + A^f,$$

where A^c and A^f are given by equations (22) and (23), and K^c and K^f are determined by the no arbitrage condition.

- The aggregate transition functions, G and H, are consistent with individual decisions as detailed in online Appendix B.

3 Individual decisions with exogenous prices

In this section, we analyze individual savings decisions when prices are constant and exogenous. We also characterize the optimal social security tax for each agent type, assuming that the social security tax remains constant. At this point, we do not conjecture how these preferences are aggregated within a political process. We simply illustrate how a social security system emerges in the sense that a majority of agents change their preferred tax from zero to something strictly positive, as the fundamentals of our environment (TFP levels in the city versus the farm and survival probabilities) change.[19]

Since all old agents die at the end of the period, they merely consume their income, and their preferred tax level is infinite. Hence, we start by characterizing the behavior of middle-aged agents.

[19] Online Appendix C contains the degree to which we can analytically order preferred tax levels by voter age and location.

3.1 Middle-Aged Agents

The preferred tax level of a middle-aged person, whether a farmer or a city worker, depends on the return to social security relative to the return to capital. The return to an extra unit of social security tax is $\frac{2}{\pi}$, while the return to an extra unit of capital is $1+r$. If the return to social security (capital) is greater than the return to capital (social security), the middle-aged agent will prefer to do all his saving via social security (capital). What is critical is that if the individual does not want to save, his optimal social security tax is zero and he will not save using either vehicle. Social security will only emerge if the agents want to save *and* the return to social security is greater than the return to capital.

We first consider the problem of a middle-aged agent who faces a tax level τ and enters the period with asset level a. Note that a consists of both the assets he saved while young and any accidental bequest he receives from his parent. Let I_m (I_o) be the total (wage and land) income of the middle-aged (old) person. An individual's age-income profile is essential in predicting his support for social security. If it is steep enough, he wants to borrow against the future instead of saving for it, and will not support social security. On the other hand, if he is better off moving resources to his old age, social security may provide an attractive way of doing so. Since a landless-middle-aged farmer will inherit land when he is old, he has a relatively steep age income profile (compared to city and landed-middle-age individuals) and is less likely to want to use social security to transfer even more resources to his old age. Proposition 1 characterizes the middle-aged agent's optimal savings decision. All the proofs are in online Appendix D.

Proposition 1. *Let a' be the optimal asset choice for a middle-aged agent. If $a' > 0$, then*

i. *if* $\frac{2}{\pi} > 1+r$, $\frac{\partial a'}{\partial \tau} < -1$;

ii. $1+r > \frac{\partial a'}{\partial a} > 0$;

iii. $1 > \frac{\partial a'}{\partial I_m} > 0$ *and* $0 > \frac{\partial a'}{\partial I_o} > -1$.

First, if the return to social security is greater than the return to capital, a middle-aged individual's preferred asset choice will fall by more than one for one with social security taxes.[20] Second, if the initial asset level of a middle-aged person rises, his optimal asset decision increases

[20] If the return to social security is less than the return to capital, no middle-aged person, and as we will show in the next section, no young person will choose a positive tax level. As a result, how the asset choice changes with the tax level is not relevant when $\frac{2}{\pi} < 1+r$.

because he will shift some resources to his old age. Third, reflecting standard life-cycle considerations, the optimal asset decision of a middle-aged person is increasing in his middle-age income and decreasing in his old-age income, so as his age-income profile flattens, he prefers to save more.

We next characterize the middle-aged individual's optimal tax decision.[21]

Proposition 2. *Let τ be the optimal tax choice for a middle-aged voter with initial asset level a. Let a' be the optimal asset choice, given τ. If $\frac{2}{\pi} < 1+r$, then $\tau = 0$. If $\frac{2}{\pi} > 1+r$, then $a' = 0$.*

Regardless of the initial asset level, every middle-aged person will choose a zero tax level if the return to social security is less than the return to capital. At his preferred tax level, a middle-aged agent will choose to save zero assets, doing all his savings via social security. However, if the prevailing tax level is not optimal for a given middle-age person, he may wish to save a strictly positive amount of assets to supplement his social security savings.

A middle-aged agent's optimal tax level depends on his initial asset level, and his income profile in a similar fashion as his optimal asset level does. Proposition 3 describes these relationships.

Proposition 3. *Let τ be the optimal tax choice for a middle-aged agent. If $\tau \geq 0$, then*

i. $1 + r > \frac{\partial \tau}{\partial a} \geq 0$;

ii. $1 > \frac{\partial \tau}{\partial I_m} \geq 0$ *and* $0 \geq \frac{\partial \tau}{\partial I_o} > -1$.

The preferred tax level of a middle-aged agent is increasing in his initial asset level, but by less than $1 + r$, the return to those assets. As income in middle age, I_m, rises, his preferred tax level rises, and as income in old age, I_o, rises, his preferred tax level falls. As a result, the landless-middle-age farmers, who have a steep age-income profile, are less likely to support social security.

3.2 Young Agents

We next characterize the outcome of the optimal tax problem for young agents. This is more complicated than characterizing the problem facing the middle-aged agents for two reasons. First,

[21] We assume that an individual can only choose a non-negative tax level. We discuss the possibility of negative social security taxes in online Appendix E.

there is an extra period which gives assets an alternative role of consumption smoothing across periods that social security cannot fill. The return when old to saving one unit of consumption when young and middle-aged via social security is $\frac{2}{\pi}$. The return when old to saving one unit of consumption when young and middle-aged via assets is $(1+r)^2+1+r$. We refer to this as private saving that mimics social security. A young agent compares $\frac{2}{\pi}$ with $(1+r)^2+1+r$ when determining if he wants a positive level of social security.

Second, young agents anticipate accidental capital bequests from their middle-aged parents, and these depend on the tax level. We show in Proposition 1 that as the tax level rises, the asset level of the middle-aged person falls (assuming the asset level is positive, otherwise it is constant at zero). When choosing his optimal tax level, the young person takes into account the non-positive effect the tax has on his potential bequest. As a result, from a young person's perspective, the total return in old age of an extra unit of τ is given by $\frac{2}{\pi}+(1+r)^2\frac{\partial b}{\partial \tau}$, where $\frac{\partial b}{\partial \tau}\leq 0$, and the growth factor $(1+r)^2$ reflects the fact that the agent will consume his bequest when he is old. For a young person who gets no bequest, the total return in old age of an extra unit of τ, is given by $\frac{2}{\pi}$, as $\frac{\partial b}{\partial \tau}=0$. Note that the cost of social security is higher for middle-aged agents whose parents survive as they experience a lower income in their middle-age.

With accidental bequests, there is uncertainty about what a young agent's asset level will be when entering middle-age. Therefore, the young individual chooses an asset level in each state of his middle-aged world: when his parent dies and he receives a bequest, and when his parent survives and he receives no bequest. Let a_1 be the asset decision while young, a_2^0 be the asset decision when middle-aged without a bequest, and a_2^1 be the asset decision when middle-aged with a bequest.

In Proposition 4, we characterize the young person's optimal tax decision.

Proposition 4. *Let τ be the optimal tax choice for a young agent. Let a_1 be the optimal asset choice while young, a_2^0 be the optimal asset choice while middle-aged without a bequest, and a_2^1 be the optimal asset choice while middle-aged with a bequest, given τ. If $\frac{2}{\pi}+(1+r)^2\left.\frac{\partial b}{\partial \tau}\right|_{\tau=0}<1+r+(1+r)^2$, then $\tau=0$. If $\frac{2}{\pi}+(1+r)^2\frac{\partial b}{\partial \tau}>1+r+(1+r)^2$, then $\tau\geq 0$, and it follows that if $a_1>0$ and $a_2^0>0$, then $a_2^1=0$; if $a_1>0$ and $a_2^1>0$, then $a_2^0=0$; and if $a_2^0>0$ and $a_2^1>0$, then $a_1=0$.*

If the return to private saving that mimics social security is greater than the return to taxation, then the young person will choose a zero tax level. Because the effect of the tax level on the bequest is non-positive, this part of the proposition implies that if $\frac{2}{\pi}<1+r+(1+r)^2$, all young people will prefer a zero tax level. On the other hand, if the return to taxation, including the non-positive effect it has on his potential bequest, is greater than the return to private saving that mimics social security, the young person may choose a positive tax level, but he will never choose a strictly positive asset level in all states. At least one of his asset level choices must be zero. The agent still may choose to use assets to smooth between periods one and two, or between periods two and three, but he will not choose strictly positive savings in all states. This simply reflects the fact that social security, which yields a higher return, provides a better way to move resources to old age, hence having positive savings in all states is not optimal. Once again, as in the case of middle-aged individuals, if the young agent is choosing the tax level, he will only choose an operative social security system if its return is greater than the return to private saving that mimics social security, and if he wants to save for old age in each state of both of the first two periods.

3.3 How Social Security Emerges

Suppose everyone lives on the farm, $\lambda = 1$, and that the return to capital is such that

$$1 + r < \frac{2}{\pi} < 1 + r + (1 + r)^2.$$

In this situation there is measure one of young farmers, who all prefer no social security (Proposition 4), measure π of middle-aged-landless farmers and measure $1 - \pi$ of middle-aged-landed farmers, who prefer a positive social security tax *if* they want to save (Proposition 2), and measure π of old farmers, who would like an infinite social security tax. The total population is $2 + \pi$. What is the preferred tax level of the middle-aged-landless farmer? Since $1 + r < \frac{2}{\pi}$, we know from Proposition 2 that if he wants to save, he will save entirely through social security, and his asset choice is zero. His preferred tax level satisfies the following first order condition

$$-u'\left(w^f \varepsilon_m^f + (1 + r)a - \tau\right) + 2\beta u'\left(w^f \varepsilon_o^f + q + \frac{2\tau}{\pi}\right) \leq 0.$$

If $2\beta u'\left(w^f \varepsilon_o^f + q\right) < u'\left(w^f \varepsilon_m^f + (1 + r)a\right)$, then the agent would like to consume more when he is middle-aged, and his optimal choice is $\tau = 0$. If $\beta \leq \frac{1}{2}$, this implies that if income in old-age $(w^f \varepsilon_o^f + q)$ is greater than income in middle-age $(w^f \varepsilon_m^f + (1 + r)a)$, the preferred tax level is zero. In the more relevant case, when $\beta > \frac{1}{2}$, if we assume a constant elasticity of substitution (CES) utility function with curvature parameter σ, then $w^f[\varepsilon_o^f - (2\beta)^{1/\sigma}\varepsilon_m^f] + q - (2\beta)^{1/\sigma}[1 + r]a > 0$, is sufficient for $\tau = 0$. The wedge in income between old-age and middle-age for a landless farmer is given by the difference in labor earnings between the two ages plus the land earnings he receives in old age minus his discounted asset income in middle-age. If this wedge is big enough, then the middle-aged-landless farmer will prefer a zero tax level. If the middle-aged-landless farmers prefer a zero social security tax, then they along with the young make up a majority of the population, $\frac{1+\pi}{2+\pi} > \frac{1}{2}$, who oppose a social security system. What is critical here is that even when social security provides a high return, i.e. $2/\pi > 1+r$, the middle-aged-landless farmers will choose not to implement social security if they do not want to shift resources to old-age.

Suppose a change in the model environment leads to rural-urban migration due to a relative increase in city wages or higher survival probabilities.[22] Once migration takes place, there is measure one of young, regardless of location. Given the assumption that $\frac{2}{\pi} < 1 + r + (1+r)^2$, they all prefer a zero tax level. There is measure π of the old, all of whom prefer an infinite amount of social security. However, through migration the distribution of middle-aged agents is now spread across landless farmers, $\lambda\pi$, landed farmers, $\lambda(1-\pi)$, and city workers, $1-\lambda$, who may or may not prefer a positive level of social security. Since $\frac{2}{\pi} > 1+r$, the optimal tax level of the landed farmers and city workers, who both have relatively flat earnings profiles, satisfies the following first order condition

$$-u'(I_m + (1+r)a - \tau) + 2\beta u'(I_o + \frac{2\tau}{\pi}) \leq 0.$$

Given that marginal utility of consumption is decreasing, if $\beta > \frac{1}{2}$ and the efficiency units are constant or decreasing with age in each location, $\varepsilon_o^j \leq \varepsilon_m^j$, where $j = c, f$, then the tax level that satisfies this inequality must be strictly positive. Therefore, these individuals prefer an operative social security system as long as $\beta > \frac{1}{2}$ and there is not an increase in labor efficiency between middle and old-age. It is also possible that through price movements induced by the technological change and migration that landless farmers also now prefer a positive tax.[23] However, even if they do not, through migration (as λ declines), the mass of individuals who support a social security system, $\lambda(1-\pi)+(1-\lambda)+\pi$, increases, while the mass of those who oppose it, $1+\lambda\pi$, decreases. In this example, if the landless farmers remain opposed, once $\lambda < \frac{1}{2}$, a majority prefers a positive

[22]Indeed, the effects of an increase in longevity are not obvious. Land is a fixed factor on the farm, so increasing survival probabilities reduces farm wages, but also increases the return to land. This crowding of land could encourage young farmers to migrate to the city. With higher life expectancy one also waits longer to inherit land, but is more likely to survive to old age. Whether an increase in life expectancy leads to migration is a quantitative question, which we explore in Section 5.

[23]See Section 5.2.2 for an example of when price movements cause middle-aged-landless farmers to support social security.

tax. Middle-aged-city workers and middle-aged-landed farmers would like to implement social security even though their age-income profiles are relatively flat. This is because middle-aged agents only pay into the system one period, while their benefits are based on two periods of payments, making their return to social security high.

3.4 Caveats

We have so far made several strong modeling choices and as a result a few caveats are in order. First, it is possible to relax (as we do in our quantitative analysis) the assumption that young farmers who migrate lose all claims to land. As long as land income declines with the structural transformation and does not provide a significant source of income, the landless will be in favor of social security.

Second, we restrict our attention to a pay-as-you-go system. This is an unrealistic description of the system in its initial years. During the first ten years, expenditures were about 15% of total receipts and social security assets grew substantially. An alternative strategy would be to allow agents to vote over both the level of benefits and the tax level. Leaving aside the political economy complications of voting over two issues, it is not clear adding this dimension to our environment would change our conclusions. Agents in our model are not altruistic and there is no aggregate uncertainty. As a result, a middle-aged median voter does not have any incentive to propose a higher tax than the one necessary to finance the system. Furthermore, by 1957, twenty years after its introduction, the system resembled a pay-as-you-go system with a ratio of expenditures to receipts of about one (Social Security Administration 2012, Table 4.A1).

Third, we abstract from exogenous population growth or endogenous fertility decisions. The fertility rate (measured as the children ever born or the total fertility rate) declined from about 5 children in the early 1800s to about 2.5 children by the early 1900s (Jones and Tertilt 2008). As a result, the population became older over these years: among the adult (ages 15 years or older) population, the fraction of 15 to 24, 25 to 44, 45 to 64 and more than 65 years old were 33.2%, 42.2%, 19.6% and 4.9% in 1870. These fractions were 29.9%, 42.7%, 20.9%, 6.2% in 1900 and 24.2%, 40.2%, 26.4% and 9.1% in 1940 (Bureau of the Census 1949, Table B 81-144). It is likely that the aging of the population increased the support for social security. By abstracting from the demographic changes of the population, we stack the cards against the emergence of social security. On the other hand, abstracting from the population structure, as we show in our quantitative exercises, gives a relatively more important role to the technological change vis a vis demographics.[24]

Fourth, we assume that both capital and social security are safe assets. It is possible to argue that compared to government bonds or the stock market, social security is a safer investment option. This would provide an additional motive for introducing social security and possibly allow us to link its emergence to the Great Depression. Although, as we mentioned in the Introduction, the role the Great Depression played as an impetus for social security is not clear.

Finally, since we abstract from population growth, it is also quite natural to treat land as a fixed factor in the model. Although large tracks of land were available in the West, the main impetus for East-West migration was the population pressure in the East. Indeed, creating productive land was quite expensive and these costs were only incurred due to population pressure (see Vandenbroucke, 2008).

4 Political Economy

To this point, we have assumed the tax level is constant and taken as given by individuals. We demonstrated qualitatively how an agent determines his own optimal tax in this setting. In order to evaluate the model quantitatively, we now focus on how the social security system is determined by equilibrium voting of successive generations. It is not obvious whether an equilibrium with social security can be supported as a political outcome in a democratic voting process with non-altruistic agents. The current young and middle-aged do not benefit from the system, yet their support is critical. Indeed, the current young and middle-aged will always choose to pay nothing in the current period, as long as they believe that the system will be there for them in the future.

We consider a variant of *constant* social security taxes: (i) if a social security system is not in place, it may start at any point, (ii) once a system is operating the tax remains constant, and successive generations take a simple yes/no vote whether or not to keep the existing system. In order to induce the agents in this economy to vote for social security according to this simple rule, we introduce a *reputational* mechanism: if a majority of voters deviate from the social security system, then the system cannot be implemented *next* period. Given that agents live for three periods, we only need to punish a deviant (middle-aged) median voter for one period.

In particular, we assume that as long as successive generations of median voters prefer a

[24] From another point of view, the introduction of the social security might lower fertility as parents are less likely to have children to care for their old age – Boldrin, De Nardi and Jones (2005).

zero tax level, then the social security system does not start. If at some point, the preferred tax level of the median voter is positive, then this tax is proposed by the median voter for a yes/no vote, and, with single-peaked preferences, is accepted by the majority as the current tax level. This tax level then remains in effect as long as successive generations of median voters prefer to keep it rather than get rid of it and live without social security. The median voter that proposes a positive tax level for the first time chooses a tax level that maximizes his lifetime utility (assuming rationally that this tax level will remain forever). If this median voter prefers a positive social security tax but instead proposes a zero tax (hoping that the next generation will implement it and he will simply benefit from it), the system cannot start next period. This ensures that median voters do not procrastinate. As a result of this mechanism, when a median voter sees that the social security tax is zero, there can be two scenarios: i) the system was never implemented and can start this period, ii) the last period's median voter deviated and chose a zero tax, and the system cannot start today. As a result, and as we detail in online Appendix F, this reputational mechanism requires that the median voter knows whether the last period's median voter deviated (i.e. did not start a system that was optimal for him) or not.[25]

The reputational mechanism we impose requires that the current voters not only know the last period's social security tax level, but also whether the last period's median voter deviated. Although this is a rather restrictive assumption, it is required to determine if the last period's median voter was procrastinating. The current median voter needs to know if the last period's median voter did not start the system with the hope that social security would start this period. The only way to punish this type of procrastination is to know whether the last period's median voter deviated.

Obviously one can consider other ways to aggregate agents's preferences. Suppose we simply assume that when a majority of citizens are better of with social security, a benevolent government starts the system and it continues as long as a majority of the citizens are better off with the system. The quantitative results with such a system would be identical to the ones presented below as long as the government chooses a tax level that maximizes the lifetime utility of the median voter. There are of course many levels of social security taxation that a majority of

[25]The reputational mechanism we use follows the standard political economy approach in the literature (see Cooley and Soares 1999, Galasso 1999, and Boldrin and Rustichini 2000). Two early papers that emphasized the political sustainability of social security were Browning (1975) and Sjoblom (1985). In the current analysis we focus on taxes that maximize the lifetime utility of the median voter, although there can be many constant tax levels that are sustainable under the reputational mechanism we have just described.

agents might prefer to not having any social security taxes and the government can implement any one of them. The political economy mechanism we describe simply allows us to pick one such tax level.

5 Quantitative Examples

Consider now the general equilibrium framework from Section 2. Although the basic intuition from the analytical results of Section 3 remain valid, there are now general equilibrium effects at play as well. This is critical for two reasons. First, the changes in relative productivity levels and survival probabilities will not only determine farm wages and land returns via migration, but will also affect all prices via changes in individual capital accumulation decisions. Therefore, it is fundamentally a quantitative question if the exogenous forces we consider and the general equilibrium effects that follow can generate a farm-to-city migration that is consistent with the data. Second, in their decisions about the social security system, agents compare the return to capital with the return to social security, but the return to capital is an endogenous variable. This is important because while higher TFP levels after 1800 push the interest rate up, higher capital accumulation associated with longer lives pushes it down. Since, as we have emphasized above, the interest rate plays an important role in the optimal tax choice, general equilibrium effects on the interest rate are of fundamental importance to the question at hand.

We now show that a calibrated version of this economy generates an initial steady state in which a majority of the population lives on the farm and the median voter chooses not to introduce a social security system, and a new steady state in which the median voter chooses a positive and sustainable social security tax. In online Appendix G, we show the transition from the initial to final steady state. We interpret the initial steady state as the U.S. economy in 1800 and the final steady state as the U.S. economy in 1940. In order to develop quantitative implications of this model economy, we first choose functional forms for utility and production functions and assign parameter values.

Let the utility function be $u(c) = \log(c)$. Since the production side of our model economy closely follows Hansen and Prescott (2002), we borrow both functional forms and parameter values from them. In particular, we assume that the production function on the farm sector is given by

$$Y^f = \gamma^f \left[N^f\right]^{\mu} \left[K^f\right]^{\phi} [L]^{1-\mu-\phi},$$

and in the city sector it is

$$Y^c = \gamma^c \left[N^c\right]^{1-\theta} \left[K^c\right]^{\theta}.$$

These choices imply that

$$w^c = (1-\theta)\gamma^c(N^c)^{-\theta}(K^c)^{\theta}, \tag{12}$$

$$w^f = \mu\gamma^f(N^f)^{\mu-1}(K^f)^{\phi}, \tag{13}$$

$$q = (1-\mu-\phi)\gamma^f(N^f)^{\mu}(K^f)^{\phi}, \tag{14}$$

and

$$r = r^c = \theta\gamma^c(N^c)^{1-\theta}(K^c)^{\theta-1} - \delta = \phi\gamma^f(N^f)^{\mu}(K^f)^{\phi-1} - \delta = r^f. \tag{15}$$

The parameter values we use are $\mu = 0.6$, $\phi = 0.1$, and $\theta = 0.4$.[26] We set the length of a model period to 20 years. We also assume that capital depreciates completely, i.e. $\delta = 1$.

Next we select the values for relative TFP levels and survival probabilities. We take TFP numbers from Greenwood and Uysal (2005). For the 1800 economy we set $\gamma^f_{1800} = \gamma^c_{1800} = 1$. Since the relative TFP values are the key determinants of migration decisions in the model, we keep $\gamma^f_{1940} = 1$ and set $\gamma^c_{1940} = 2.19$. These choices imply that the relative TFP growth is as reported by Greenwood and Uysal (2005) and reproduced in Figure 4. Historical estimates for age-specific-mortality rates and life tables do not go back further than 1850 (see Haines 1998). In 1850, a 60 year-old man had about a 47% chance of living to his 80th birthday. Since available evidence does not indicate any significant improvement in mortality between 1800 and 1850, we set $\pi_{1800} = 0.47$.[27] In 1940 the chances that a 60 year old man saw his 80th birthday increased to about 58%. Therefore, we select $\pi_{1940} = 0.58$.[28]

[26] The value for capital share in the city (industrial) technology, $\theta = 0.4$, is the standard value for the postwar U.S. economy. The labor share is assumed to be the same for both sectors, $\mu = 1 - \theta = 0.6$. Finally, $\phi = 0.1$ is picked to be consistent with historical evidence on agricultural incomes. See Hansen and Prescott (2002) for details.

[27] According to Haines (1998), the crude death rate in New York City was as high in 1850 as it was in 1804 (see Figure 1, page 150). In many New England towns there was not much improvement in life expectancy at age 20 either (see Table 1, page 151).

[28] Let P_i be the size of age-i population. We have data on P_{60-64}, P_{65-69}, P_{70-74}, P_{75-80} and P_{80+} in 1850, and on $P_{60}, P_{65}, .., P_{80}$ in 1940. For 1850, we calculate π as

$$\pi = \frac{(P_{60-64} + P_{65-69} + P_{70-74} + P_{75-80} + P_{80+})/4}{P_{60-64}},$$

which captures an average survival rate. The calculations for 1940 are done in a similar way.

Finally, we assume that agents have flat age-earning profiles both on the farm and in the city, i.e. $\varepsilon_i^j = 1$ for $j \in \{f, c\}$ and $i \in \{y, m, o\}$. Age-earning profiles in the 19th century did indeed differ from the usual hump-shaped pattern. According to Kaelble and Thomas (1991), incomes of working class household heads increased slightly between ages 20 and 40, but were pretty much flat after age 40. These flat profiles were a common feature of agricultural workers as well as low skilled non-agricultural workers.[29] We make the strong assumption that age-earning profiles were also flat in the city. We consider this to be a conservative assumption for this exercise, since a hump-shaped profile for city workers would simply increase the incentives of middle aged workers to shift resources to their old age and increase the political support for social security even further.

Note that all of these parameter values are fixed prior to running our simulations. We are left with only one more parameter to pick, β. We set $\beta = 0.818$ (a yearly value of 0.99). This value implies that the yearly return to capital in the 1940 steady state is about 6.1%.[30] Table II summarizes our parameter choices.

Table II — Parameter Values

β	μ	ϕ	θ	δ	γ_{1800}^f	γ_{1800}^c	γ_{1940}^f	γ_{1940}^c	π_{1800}	π_{1940}
0.818	0.6	0.1	0.4	1	1	1	1	2.19	0.47	0.58

5.1 Results

Table III shows the results. In our 1800 economy everyone lives on the farm, $\lambda = 1$. This is consistent with the U.S. experience. At that time, about 94% of population lived in rural areas, and the fraction of population working on the farm was possibly even higher (see Figure 1). In the 1800 steady state, the median voter is a landless-middle-aged farmer, who does not want social security, so the equilibrium value of τ is zero. Notice that this happens even though $2/\pi$ (about 4.25) is larger than $1 + r$, so the direct return to social security is greater than the

[29] Doepke and Zilibotti (2008) contrast relatively flat wage profiles of agricultural workers and land owners with steep wage profiles of entrepreneurs in the 19th century. They model the emergence of capitalism within a model of structural transformation in which entrepreneurs influence their children's preferences in an attempt to make them more patient.

[30] Cooley and Prescott (1995) report a value of 6.9 percent rate of return on capital for post-war period. See Gomme and Rupert (2007) for a more recent discussion.

return to capital. However, the middle-aged-landless farmer prefers to save nothing due to his steep age-income profile.

Next, consider the 1940 economy. Now about 23% of the population lives on the farm, which is exactly what is observed in the U.S. at that time (see Figure 1). This is quite remarkable since nothing in our parameter choices targets directly the fraction of agents living on the farm. With migration to the city, the median voter is a city worker whose preferred tax level is 0.106, about 19% of city wages.[31]

Table III - Initial and Final Steady States

	1800	1940
τ	0	0.105
λ_y	1	0.227
$1+r$	2.466	3.247
w^f	.311	0.487
w^c	.178	0.549
q	0.384	0.142
q/λ	0.384	0.626
K	0.052	0.240
K^f	0.052	0.012
K^c	0	0.228
N^f	2.470	0.587
N^c	0	1.997
Median Voter	middle-age-landless farmer	middle-aged city worker

Consistent with historical experience, the return on capital is much higher in the new steady state, despite an almost fourfold increase in aggregate capital stock. In 1940, about 23% of the population lives on the farm, but a much smaller (about 5%) fraction of aggregate capital stock is allocated to farm production. Also consistent with historical evidence, the rental value of land declines significantly. In 1940 it is about 37% of its 1800 value.[32] Lastly, note that while the

[31] When social security was introduced the total (employee plus employer) tax rate was about 2%, which has gradually increased to its current level of 15.3%.

[32] According to Hansen and Prescott (2002), the value of U.S. farmland relative to GDP declined from 88% in 1870 to 20% in 1950 (see Table 2, page 1209).

returns per unit of land, q, fall, the returns to land for landholders, q/λ, actually rise, 0.38 to 0.63, which keeps people on the farm despite rising city wages.

In the new steady state, even though total labor supply in the city rises due to the increases in life expectancy and migration, because of the increases in technological progress and in the aggregate capital stock, the city wage rises. There is no technological advance on the farm. But the out migration of farmers causes labor supply on the farm to fall, and so farm wages rise.

It is worth noting that the demographic changes alone would not lead to the rural-urban transition that the U.S. experienced. When we only change survival probabilities, social security does not emerge, because the change does not induce enough migration. Indeed, everybody remains on the farm. The key effect of this change is an increase in the capital stock because people save more anticipating a longer life.

When only TFP changes, social security does emerge as an equilibrium outcome but the rural/urban migration is not nearly as pronounced. Roughly 30% continue to live on the farm (in the data and in our economy with changes in both survival probabilities and the TFP it is 23%). Furthermore, the social security tax is higher than in the economy with both factors at work. This underscores the conclusion that neither technology nor demographics alone is sufficient to account for events but the interaction between the two is a powerful impetus for social change.

5.2 Extensions

In this section we consider three extensions of our basic model. First, we allow agents to insure against mortality risk. Annuity markets are quite thin today and they were certainly absent in the 1940s (as well as 1800).[33] Almost all quantitative work on social security assumes that they are not available to agents.[34] As a result, although annuities can (and do) play a role theoretically, the assumption of a well-functioning annuities market is at odds with historical experience. Introducing annuities to our framework crowds out the insurance role that social security is currently providing. As one might expect, this reduces support for social security. In an environment with perfect annuities, the level of social security that emerges is rather small. Although, in a world with partial annuities (which is a more realistic assumption), we see social security arising to an extent similar to that in our benchmark final steady state. Allowing for

[33] Only 2 to 4% of elderly owned private annuities between the 1930s and the 1980s according to Warshawsky (1988).

[34] See the large literature starting with Imrohoroglu, Imrohoroglu and Joines (1985).

annuities eliminates accidental bequests of capital and highlights the role played by land (the main form of inheritance in the 19th century) in the initial steady state, which we assume cannot be annuitized.

Second, we allow agents who migrate to the city to claim income from the land they inherit from their parents. It may seem restrictive that we deny migrants their entire land inheritance. At first glance, it appears that this assumption is driving the result that social security emerges in the final steady state. This is because as farmers migrate, if they keep their land inheritance, there will be middle-age-landed and -landless city workers. The landless-city workers face an age-income profile that is similar to the landless farmers, in the sense that their future land income in old age creates a positive wedge in income across periods. It turns out that through migration this wedge shrinks to the point where it no longer causes the middle-age landless (in the city or on the farm) to oppose social security.

Lastly, we consider what happens when farmers are exempt from the social security system. When the social security system was introduced farmers and other self-employed workers were not covered by the system. They remained out of the system until the 1950 amendments (see DeWitt 2010). In this case, a farmer's preferred tax level only depends on its resulting general equilibrium effects. We find that a farmer's utility as a function of the tax level is not necessarily single-peaked, so we limit voters to two choices: no social security and some positive level of social security. For a very small city population, we find the median voter still chooses a zero tax level. Once TFP and survival probability increase to their 1940 levels, a positive level of social security can be sustained.

5.2.1 Annuities

Incorporating annuities into the current framework alters individual savings decisions. In particular, it shifts up the return to assets relative to social security. As our analytical results show, when middle-aged agents choose whether or not to support social security they compare the return to social security, $\frac{2}{\pi}$, with the return to assets, $1 + r$. With annuities, the return to assets rises to $\frac{1+r}{\pi}$, so the comparison is now 2 vs. $1 + r$. One role social security plays in our original framework is to provide insurance against mortality risk. Therefore, introducing annuities undermines support for social security. If the median voter wants to save, he is more likely to choose assets over social security.

However, this is only part of the story as there are also general equilibrium effects resulting from the implementation of social security. In particular, the capital stock falls with the introduction of the social security. This implies an increase in the return to capital and a decrease in wages compared to a world without social security. Hence, the median voter in 1940, a city worker, might prefer to have social security (even if it provides a relatively low return and lowers his wages) due to the higher returns on his savings it implies. The first two columns of Table IV contain the results with perfect annuities. Social security emerges in 1940 as the median voter prefers social security as a result of its general equilibrium effects on the returns to his savings. The preferred tax level is much lower than when there are no annuities, highlighting the fact that the return to assets is now higher. Note that there is more capital now as compared to the steady states with no annuities (Table III), reflecting the higher return on savings. This higher level of capital also increases city wages (compared to a world without annuities), which contributes to a higher level of migration.

Table IV- Extensions

	Perfect Annuities		Imperfect Annuities		Migrants Inherit Land	
	1800	1940	1800	1940	1800	1940
τ	0	0.014	0	0.098	0	0.090
λ_y	1	0.165	1	0.240	1	0.154
$1+r$	1.961	2.412	2.397	3.145	2.466	3.193
w^f	0.319	0.560	0.315	0.480	0.311	0.555
w^c	0.208	0.672	0.192	0.560	0.178	0.555
q	0.394	0.119	0.389	0.149	0.384	0.111
q/λ	0.394	0.721	0.389	0.384	-	-
K	0.067	0.420	0.059	0.249	0.052	0.265
Median Voter	middle-age landless farmer	middle-age city worker	middle-age landless farmer	middle-age city worker	middle-age landless farmer	middle-age city worker

If annuities are not perfect, a higher level of social security emerges as a result of migration. Suppose, for example, that the return to annuities is $\frac{1+r}{\pi(1+\chi)}$ instead of $\frac{1+r}{\pi}$, with the parameter χ reflecting the percentage leakage in the annuity system.[35] As long as $1+\chi < \frac{1}{\pi}$, agents still get an asset return that is higher than they would get without annuities. The second two columns of Table IV show the results with $\chi = 0.5$. With this level of χ, we have a final steady state that looks similar to the final steady state in Table III.

[35] In order to keep the computational analysis simple we assume that this leak is returns that are not transferred to farmers (not a leakage on actual capital). It is lost after the production take place.

5.2.2 City Workers with Land Income

Suppose young farmers can receive land inheritances even after migrating to the city. The last two columns of Table IV contain the steady state results. Note that city and farm wages are equal in the final steady state. Since migrants to the city can now earn land income, the only difference in the two sectors is the wages. If city wages were higher (as was the case in the economy where migrants lose claims on land inheritances), then everyone would move to the city. The final steady states in Table III and the last column of Table IV look fairly similar, suggesting that the landless-middle age, in this exercise, must now support social security. In our original framework, the farm land that is left behind by those who migrate is spread among the remaining farmers. So, the return to land is $\frac{q}{\lambda}$. If migrants keep their land, however, then the return is q, as there is no land left behind to redistribute. In the new steady state, the return to land, q, falls so much that it has very little effect on the city worker's age-income profile. This is an example of a situation where the changes that instigate migration, combined with the resulting migration, cause the middle-aged-landless median voter to switch from opposing social security to supporting it. Allowing city workers to collect land income turns out not to be an important restriction in our model.

5.2.3 Farmers Exempt From Social Security System

If farmers are exempt from the direct costs and benefits of social security, then their voting behavior will depend entirely on the effect the tax has on prices. As the tax increases, the capital stock will fall, pushing the return to capital up. For those in the city, this increase in return to capital is accompanied by a higher tax payment. But farmers reap the benefits of a higher return to capital, without paying a cost. On the other hand, as the tax rises, the city becomes a less attractive place for the marginal young farmer, causing farm wages fall. This effect is capped when the tax level gets high enough that there is no further migration. Due to these conflicting forces farmers can have non-single-peaked preferences. At very low levels of taxation, they benefit from out migration that increases their wages. At very high levels of taxation, the negative impact on wages via reduced migration ceases, but the gain from a higher return to capital continues.

Since we cannot guarantee that farmers have single-peaked preferences, in this section we limit voters to two choices: no social security and some positive level of social security. As we discuss in Section 4, doing this implies many possible levels of social security can be sustained by the median voter. The intention of this exercise is to demonstrate that, even if farmers are exempt social security does not emerge with low TFP in the city, low survival probability, and hence a small population in the city, but does still emerge when city TFP is high, survival probability is high, and there is a large fraction of population in the city.

The voting behavior of city voters does not change when farmers are exempt from the system: The young are worse off the higher the tax, the old are better off the higher the tax, and the middle-age have an optimal interior tax level. The farmers' choices do change. Farmers who have high asset levels are happier with higher taxes, since taxes boost returns to capital, while those who have low asset levels are happier with lower taxes since their wages are higher. An increase in city TFP and survival probability implies higher asset levels, and more farmers prefer a positive tax level. It is also the case that there are many fewer farmers when TFP in the city rises relative to farm TFP.

Given the parameters for 1800, there is no city population and no social security. Because there is no city population to respond to the taxation, there are no general equilibrium effects. All tax levels provide identical utility to all farmers. Therefore, the question of whether or not farmers support or oppose social security is vacuous. In order to make the question interesting, we increase the TFP in the city enough (from $\gamma = 1$ to $\gamma = 1.61$) to generate a small population in the city ($1 - \lambda = .03$.) We find that the median voter, who clearly is still a farmer, prefers a zero tax level to any positive tax level. With our 1940 parameters, we find that when the median voter is given a choice between a tax of zero and any positive tax level below $\tau = .14$, he prefers social security. Suppose the tax level is $\tau = .105$, which is what the median voter chooses when farmers are part of the system. At this level of taxation, when farmers are exempt, the farm population is 33%, which is higher than the 23% we find in the original 1940 steady state. This is no surprise, because when they are outside the system, young farmers prefer to stay on the farm and avoid paying taxes. Because of this difference in city and farm populations, city wages are higher and farm wages are lower when farmers are exempt from the social security system. Lastly, the return to capital is not as high when farmers are exempt (5.8% vs. 6.1%), because farmers save more given their higher income.

6 Conclusion

In this paper we develop a model economy in which the structural transformation from a rural to an urban economy gives rise to support for a social security system. We demonstrate how this can occur analytically within our framework. When everyone lives on the farm a majority of the population prefers to rely on land inheritance for old-age security in place of a governmental system of old-age pensions. Once enough migration takes place, a majority prefers some positive level of social security. However, those who would like to implement social security have heterogeneous preferences over how big the program should be. We assume majority voting as a method for aggregating preferences, and then quantitatively investigate whether our framework is consistent with the historical experience of the United States before and after social security is implemented. We show that there is an initial steady state consistent with the United States in the 1800s, with most people living on the farm and no social security system. Changes in life expectancy and technological progress in the city that are in line with those observed in the data give rise to a new steady state. In this final steady state, the majority of the population lives in the city and the median voter supports implementing a social security system.

One key element of our story is that middle-aged-landless farmers are unable to borrow against the land inheritance they receive in old age, conditional on survival. This borrowing constraint yields a steep age-income profile for the landless-middle-age farmers, which implies that these individuals do not want to save, even if social security provides a high return. The evidence shows that the sort of long-term borrowing required to undo this effect was not available over the time period we investigate. While alternative mechanisms may be able to provide an impetus for social security, our explanation is the first to give a quantitatively-consistent link between industrialization and demographic change, and social security.

The current framework can be used to shed light on two issues of fundamental importance. First is the question of why did different countries follow such different strategies in constructing their social safety nets, choosing different degrees of reliance on state versus the market.[36] The current model provides a natural framework to link demographics, geography, and differences in the structural transformation of countries to differences in social insurance institutions. The second is the dramatic transformation that is taking place in China. Currently, there is no national pension system (nor much in the way of social insurance) in China, but as the world's largest ever peacetime flow of migration continues, and the traditional support systems via the family are dismantled, we would expect the demand for such institutions to grow. We leave these questions for future research.

[36]Perotti and Schwienbacher (2009) study how large inflationary shocks in the first half of the XX century, which devastated middle class savings in some countries, affected their reliance on state versus market institutions.

7 Appendix A - Data Sources

Figure 1: Hernandez (1996), Table 4.

Figure 2: The urbanization rates are from the 1930 Census, Table 6, page 10, available at http://www2.census.gov/prod2/decennial/documents/16440598v2_TOC.pdf. The elderly population is calculated from Hobbs and Stoops (2002), Table 7, page A-19. The dates for state old age assistance laws are taken from ElderWed, http://www.elderweb.com/home/node/2896.

Figure 3: The fraction of the labor force in agriculture is based on Mitchell (2003), Table B1 Economically Active Population, by Major Industrial Groups, page 147. The adoption of social security dates are from the Social Security Administration (2006).

Figure 4: Greenwood and Uysal (2005), Figure 9.

Figure 5: The data for 1850 and 1900 are from Haines (1998) and for 1950 are taken from the U.S. Department of Health, Education and Welfare (1964). They are the average of the conditional survival probabilities from age 60 to 65, from 65 to 70, from 70 to 75 and 75 to 80. The 1850 numbers are for white males only.

Table I: GDP per capita is taken from Carter et al (2006), Table Ca9-19, Loans and Discounts of Banks is from Carter et al (2006), Table Cj149-157, the U.S. population, which is used to calculate per capita loans and discounts is from Carter et al (2006), Table Ca9-19. The information on maturities is from Hammond (1934), page 89.

References

[1] Amenta, Edwin and Bruce G. Carruthers. "The Formative Years of U.S. Social Spending Policies: Theories of Welfare State and the American States During the Great Depression." American Sociological Review, 53 (October 1988): 661-678.

[2] Atack, Jeremy, Bateman, Fred, and William N. Parker. "The Farm, The Farmer, and The Market," in T*he Cambridge Economic History of the United States, Vol II: The Long Nineteenth Century*, Stanley L. Engerman and Robert E. Gallman (eds.), Cambridge University Press: Cambridge, UK. 2000.

[3] Barlowe, Raleigh, and Timmons, John F. "What Has Happened to the Agricultural Ladder?" *Journal of Farm Economics*, 32 (February 1950): 30-47.

[4] Bertocchi, Graziella. "The Vanishing Bequest Tax." *Economics & Politics* 23 (March 2011): 107-131.

[5] Boldrin, Michele, and Rustichini, Aldo. "Political Equilibria with Social Security." *Review of Economic Dynamics* 3 (January 2000): 41-78.

[6] Boldrin, Michele, De Nardi, Mariacristina and Larry Jones. "Fertility and Social Security," mimeo, 2005.

[7] Browning, Edgar K. "Why the Social Insurance Budget is Too Large in a Democracy." *Economic Inquiry* 13 (September 1975): 378-88.

[8] Bureau of the Census. *Historical Statistics of the United States, 1780-1940: Supplement to the Statistical Abstract of the United States*, U, S. Government Printing Office: Washington D. C. 1949.

[9] Carter Susan B., Gartner, Scott S., Haines, Michael R., Olmstead, Alan L., Sutch, Richard and Gavin Wright. *Historical Statistics of the United States, Vol III, Part C: Economic Structure and Performance.* Cambridge University Press: Cambridge, UK. 2006.

[10] Cole, Harold L. and Ohanian, Lee E. "New Deal Policies and the Persistence of the Great Depression: A General Equilibrium Analysis." *Journal of Political Economy*, 112(August 2004): 779-816.

[11] Collier, David and Messick, Richard E. "Prerequisites versus Diffusion: Testing Alternative Explanations of Social Security Adoption." *The American Political Science Review* 69 (December 1975): 1299-1315.

[12] Conesa, Juan C., and Krueger, Dirk. "Social Security Reform with Heterogeneous Agents." *Review of Economic Dynamics* 2 (October 1999): 757-795.

[13] Cooley, Thomas F. and Prescott, Edward C. “Economic growth and business cycles” in Cooley, T. (ed.), *Frontiers of Business Cycle Research* Princeton University Press: Princeton, (1995) 1-38.

[14] Cooley, Thomas F., and Soares, Jorge. “Will Social Security Survive the Baby Boom?” *Carnegie-Rochester Conference Series on Public Policy* 45 (1996): 89-121.

[15] Cooley, Thomas F., and Soares, Jorge. “A Positive Theory of Social Security Based on Reputation.” *Journal of Political Economy* 107 (February 1999): 135-160.

[16] Corbae, Dean, D’Erasmo Pablo and Kuruscu, Burhan. “Politico Economic Consequences of Rising Wage Inequality.” Journal of Monetary Economics 56 (January 2009): 53-61.

[17] DeLong, Bradford J. “Bequests: An Historical Perspective," in Death and Dollars: The Role and Impact of Gifts and Estates, edited by Alicia Munnell and Annika Sunden (editors). Washington, DC: Brookings Institution, 2003.

[18] DeWitt, Larry. “The Development of Social Security in America.” Social Security Bulletin, 70 (August 2010): 1-26.

[19] Doepke, Matthias and Tertilt, Michele. “Women’s Liberation: What’s in it for Men?” *Quarterly Journal of Economics* 124 (November 2009): 1541-1591.

[20] Doepke, Matthias and Zilibotti, Fabrizio. “The Macroeconomics of Child Labor Regulation.” *American Economic Review* 95 (December 2005): 1492-1524.

[21] Doepke, Matthias and Zilibotti, Fabrizio. “Occupational Choice and the Spirit of Capitalism.” *Quarterly Journal of Economics* 123 (May 2008): 747–793.

[22] Echevarria, Cristina. “Changes in Sectoral Composition Associated with Economic Growth.” *International Economic Review* 38 (May 1997): 431-452.

[23] Galasso, Vincenzo. “The U.S. Social Security System: What Does Political Sustainability Imply?” *Review of Economic Dynamics* 2 (April 1999): 698-730.

[24] Galor, Oded and Moav, Omer. “Das Human-Kapital: A Theory of the Demise of the Class Structure.” The Review of Economic Studies Vol. 73 (January 2006): 85-117.

[25] Galor, Oded, Moav, Omer and Vollrath, Dietrich. “Inequality in Land Ownership, the Emergence of Human Capital Promoting Institutions and the Great Divergence.” *Review of Economic Studies* 76 (January 2009): 143-179.

[26] Galor, Oded and Weil, David N. "Population, Technology and Growth: From the Malthusian Regime to the Demographic Transition". American Economic Review 110 (September 2000): 806–828.

[27] Gollin, Douglas, Parente, Stephen, and Rogerson, Richard. "The Role of Agriculture in Development." *American Economic Review (Papers and Proceedings)* 92 (May 2002): 160-164.

[28] Gomme, Paul and Rupert, Peter. "Theory, Measurement and Calibration of Macroeconomic Models." *Journal of Monetary Economics* 54 (March 2007): 460-497.

[29] Gonzalez-Eiras, Martin and Niepelt, Dirk. "Sustaining Social Security." CESifo Working Paper No. 1494. July 2005.

[30] Greenwood, Jeremy and Seshadri, Ananth. "The U.S. Demographic Transition." *American Economic Review (Papers and Proceedings)* 92 (May 2002): 153-159.

[31] Greenwood, Jeremy and Uysal, Gokce. "New Goods and the Transition to a New Economy." *Journal of Economic Growth* 10 (June 2005): 99-134.

[32] Haines Michael R. "Estimated Life Tables for the United States, 1850-1910." *Historical Methods*, vol. 31 (Fall 1998): 149-169.

[33] Hammond, Bray. "Long and Short Term Credit in Early American Banking," *Quarterly Journal of Economics*, 49 (November 1934): 79-103.

[34] Hansen, Gary D., and Prescott, Edward C. "Malthus to Solow." *American Economic Review* 92 (September 2002): 1205–1217.

[35] Hassler, John, Rodríguez-Mora, José V., Storesletten, Kjetil and Zilibotti, Fabrizio. "Survival of the Welfare State." *American Economic Review* 93 (March 2003): 87-112..

[36] Hassler, John, Rodríguez-Mora, José V., Storesletten, Kjetil and Zilibotti, Fabrizio. "A Positive Theory of Geographical Mobility and Social Insurance." *International Economic Review* 46 (2005): 263-303.

[37] Hernandez, Donald J. *Population Change and the Family Environment of Children. Part 2 of Trends in The Well-Being of America's Children and Youth: 1996.* U.S. Department of Health and Human Services, Washington, D.C. 1996.

[38] Hobbs, Frank and Stoops, Nicole. *Demographic Trends in the 20th Century.* U.S. Census Bureau, Census 2000 Special Reports, Series CENSR-4. Washington, DC: U.S. Government Printing Office, 2002.

[39] Imrohoroglu, Ayse, Imrohoroglu, Selo, and Joines, Douglas. “A Life Cycle Analysis of Social Security.” *Economic Theory* 6, 1985 83-114.

[40] Jones, Larry and Tertilt, Michele. “An Economic History of Fertility in the U.S.: 1826-1960,” in *Frontiers of Family Economics*, edited by Peter Rupert, Elsevier, 2008.

[41] Kaelble, Hartmut and Thomas, Mark. “Introduction” In *Income Distribution in Historical Perspective* edited by Y. S. Brenner, Hartmut Kaelble, and mark Thomas. Cambridge, UK: Cambridge University Press 1991.

[42] Kim, Kyo-seong. “Determinants of the Timing of Social Insurance Legislation Among 18 OECD Countries,” *International Journal of Social Welfare* 10 (January 2001): 2-13.

[43] Klein, Paul and Ventura, Gustavo. "Productivity Differences and the Dynamic Effects of Labor Movements,” Journal of Monetary Economics 56 (November 2009): 1059-1073.

[44] Kongsamut, Piyabha, Rebelo, Sergio, and Xie, Danyang.“Beyond Balanced Growth.” *Review of Economic Studies*, 68 (October 2001): 869-882.

[45] Kopecky, Karen A. “The Trend in Retirement.” *International Economic Review*, 52 (May 2011): 287-316.

[46] Krueger, Dirk and Kubler, Felix. “Pareto Improving Social Security Reform When Financial Markets are Incomplete,” *American Economic Review* 93 (June 2005): 737-755.

[47] Krusell, Per, Vincenzo Quadrini, and Jose-Victor Rios-Rull. “Politico-Equilibrium and Economic Growth.” *Journal of Economic Dynamics and Control* 21 (January 1997): 243-72.

[48] Krusell, Per and Jose-Victor Rios-Rull. “On the Size of U.S. Government: Political Economy n the Neoclassical growth Model.” *American Economic Review* 89 (December 1999): 1156-1181.

[49] Laitner, John. “Structural Change and Economic Growth.” *Review of Economic Studies,* 67 (July 2000): 545-561.

[50] Maddison, Angus. *The World Economy: A Millennial Perspective*. Paris: OECD Publications, 2001.

[51] Miron, Jeffrey A. and Weil, David N. “The Genesis and Evolution of Social Security.” In *The Defining Moment: The Great Depression and The American Economy in the Twentieth Century*, edited by Michael D. Bordo, Claudia Goldin and Eugene N. White. Chicago, IL: The University of Chicago Press, 1998.

[52] Mitchell, B.R. *International Historical Statistics: Europe* , 1750-2000. London, UK : Palgrave MacMillan, 2003.

[53] Newell, William H. "Inheritance on the Maturing Frontier: Butler County, Ohio, 1803-1865." In *Long-Term Factors in American Economic Growth*, edited by Stanley L. Engerman and Robert E. Gallman. Chicago, IL: The University of Chicago Press, 1986.

[54] Parente, Stephen and Prescott, Edward. "A Unified Theory of the Evolution of International Income Levels." In *Handbook of Development,* P. Aghion and S. Durlauf (editors). Amsterdam, The Netherlands: Elsevier, 2005.

[55] Perotti, Enrico C. and Schwienbacher, Armin. "The Political Origin of Pension Funding." *Journal of Financial Intermediation* 18 (July 2009): 384-404

[56] Pryor, Frederic L. *Public Expenditures in Communist and Capitalist Nations*. London: George Allen and Unwin LTD., 1968.

[57] Rothenberg, Winifred B. *From Market Places to a Market Economy: The Transformation of Rural Massachusetts 1750-185*0. Chicago University Press: Chicago, 1992.

[58] Sass, Steven A. *The Promise of Private Pensions – The First Hundred Years*. Cambridge, Massachusetts: Harvard University Press, 1997.

[59] Scheiber, Sylvester J., and Shoven, John. *The Real Deal: The History and Future of Social Security.* New Haven, Connecticut: Yale University Press, 1999.

[60] Sjoblom, Kriss. "Voting for Social Security." *Public Choice* 45 (1985): 225-240.

[61] Social Security Administration. *Social Security Programs Throughout the World: Europe, 2006.* Available at http://www.ssa.gov/policy/docs/progdesc/ssptw/2006-2007/europe/index.html. Last updated 2006.

[62] Social Security Administration. *Historical Background and Development of Social Security.* Available at http://www.ssa.gov/history/briefhistory3.html. Last updated 2003.

[63] Social Security Administration. *Annual Statistical Supplement to the Social Security Bulletin, 2011.* SSA Publication No. 13-11700, 2012.

[64] Soltow, Lee. "Male Inheritance Expectations in the United States in 1870." *The Review of Economics and Statistics* 64 (1982): 252-260.

[65] Soltow, Lee. "Inequalities in the Standard of Living in the United States." In *American Economic Growth and Standards of Living Before the Civil War*, edited by Robert E. Gallman and John Joseph Wallis. Chicago, IL: The University of Chicago Press, 1992.

[66] Temin, Peter. *Lessons from the Great Depression.* Cambridge, Massachusetts: The MIT Press, 1991.

[67] U.S. Bureau of the Census. *Historical Statistics of the United States: Colonial Times to 1970.* Washington, D.C.: U.S. Bureau of the Census, 1975.

[68] U.S. Department of Health, Education and Welfare. *US Life Tables 1959-61.* Public Health Service Publication 1252, vol. 1, no.1. Washington, D.C. 1964.

[69] Vandenbroucke, Guillaume. "The U.S. Westward Expansion." *International Economic Review* 49 (February 2008): 81:110.

[70] Wang, Ta-Chen. "Banks, Credit Markets, and Early American Development: A Case Study of Entry and Competition," *The Journal of Economic History*, 68 (June 2008): 438-461.

[71] Warshawsky, Mark. "Private Annuity Markets in the United States," *Journal of Risk and Insurance* 55 (September 1988).

[72] Wiebe, Robert H. *Search for Order 1877-1920.* New York, NY: Hill and Wang, 1966.

[73] Williamson, Jeffrey G. and Lindert, Peter H. "Long-Term Trends in American Wealth Inequality." In *Modeling Distribution and Intergenerational Transmission of Wealth*, James D. Smith (editor). Chicago, IL: The University of Chicago Press, 1980.

[74] Wolf, Edward N. "Changing Inequality of Wealth," *American Economic Review (Papers and Proceedings)* 82 (May 1992): 552-558.

[75] Wright, Robert E. *Origins of Commercial Banking in America, 1750-1800.* Lanham, Maryland: Rowman and Littlefield Publishers, Inc., 2001.

[76] Yang, Donghyu. "Farm Tenancy in the Antebellum North." In *Strategic Factors in Nineteenth Century American Economic History*, Claudia Goldin and Hugh Rockoff (editors). Chicago, IL: The University of Chicago Press, 1992.

Figure 1 --- Population in Rural and Urban Areas

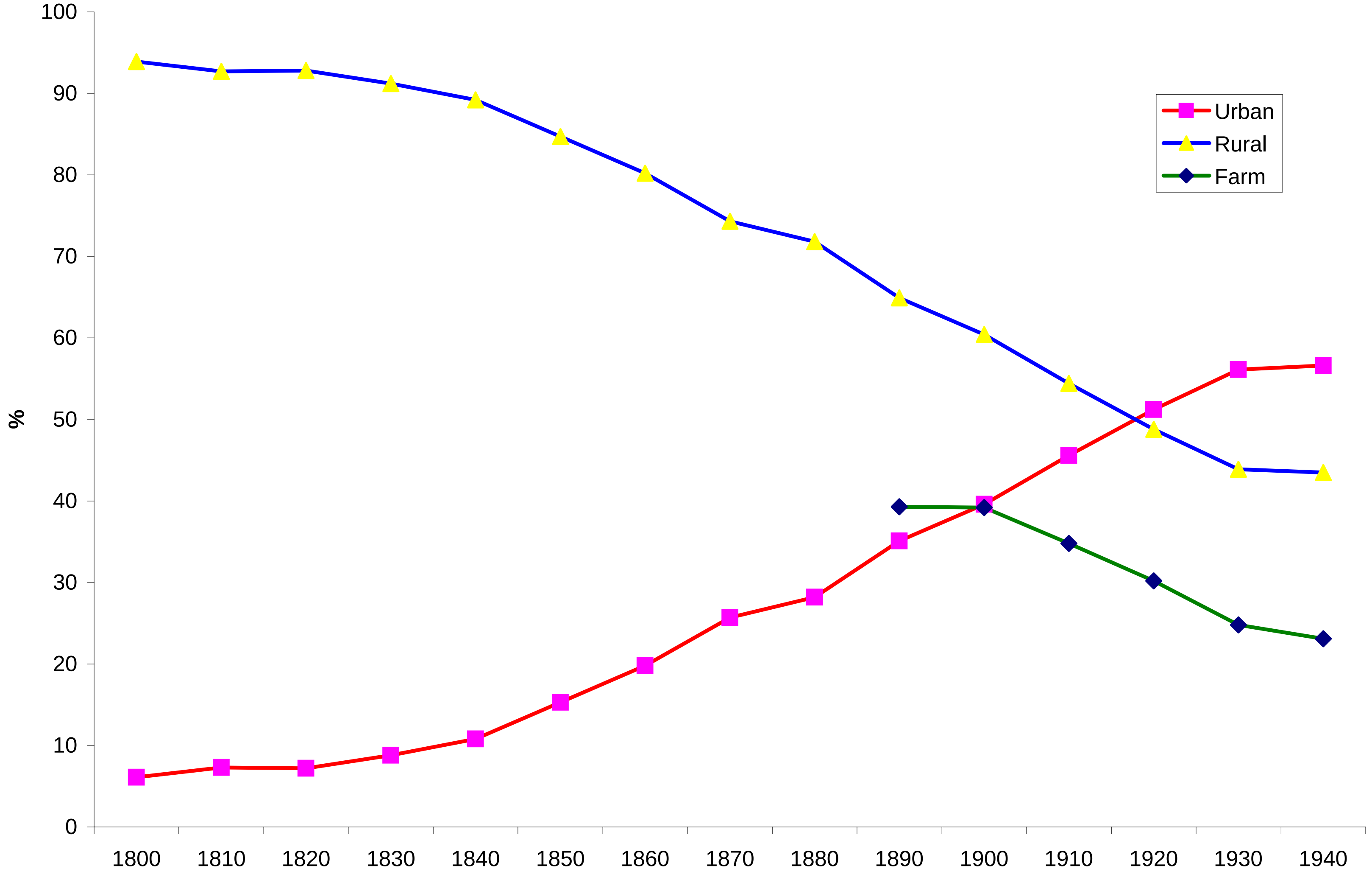


Figure 2: Urbanization and Elderly (65+) Population Across U.S. States -- 1930
States in Circles Introduce State Pension Plans before 1935

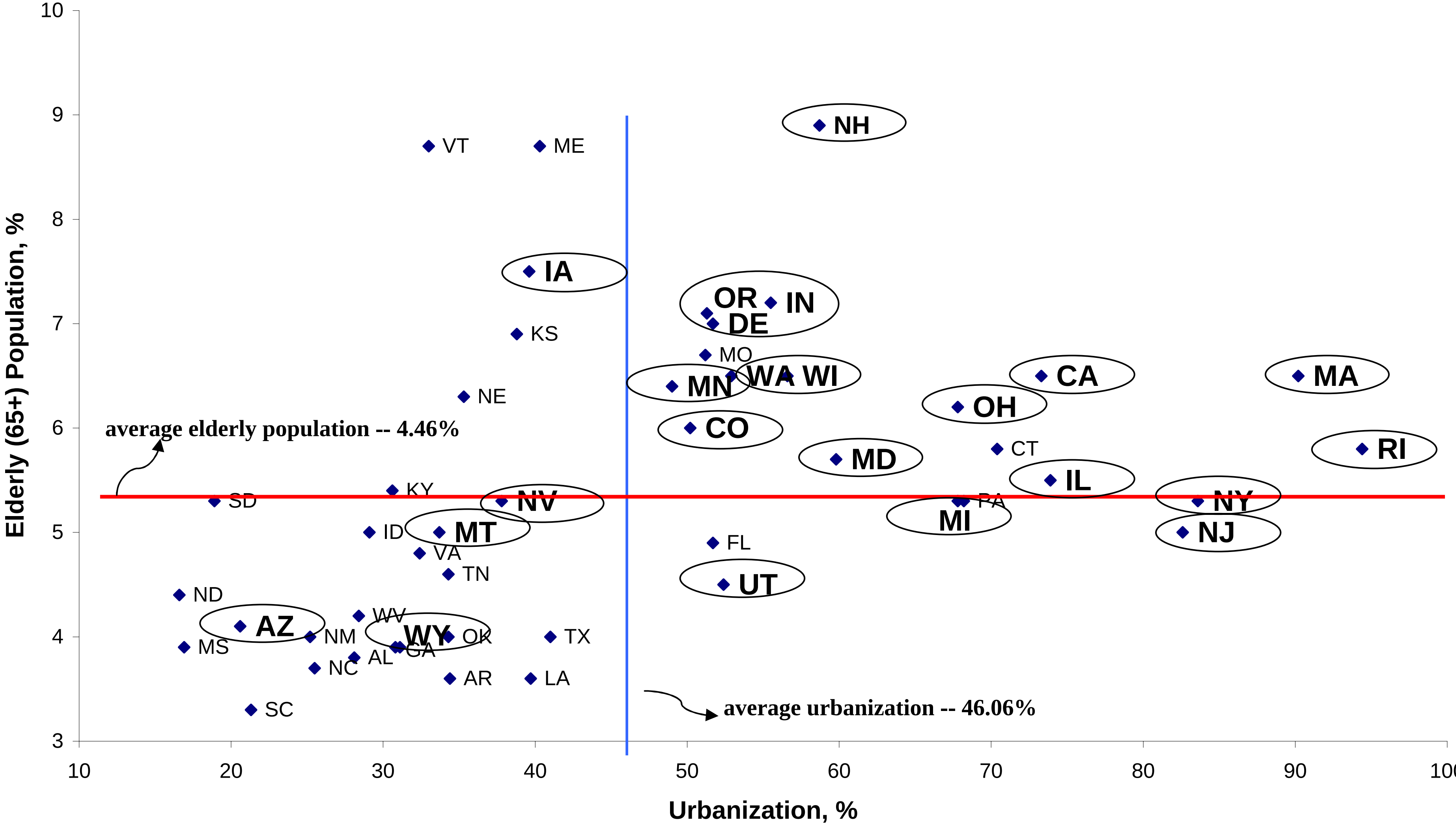


Figure 3: Fraction of Labor Force in Agriculture in 1890 and The Adoption of Social Security

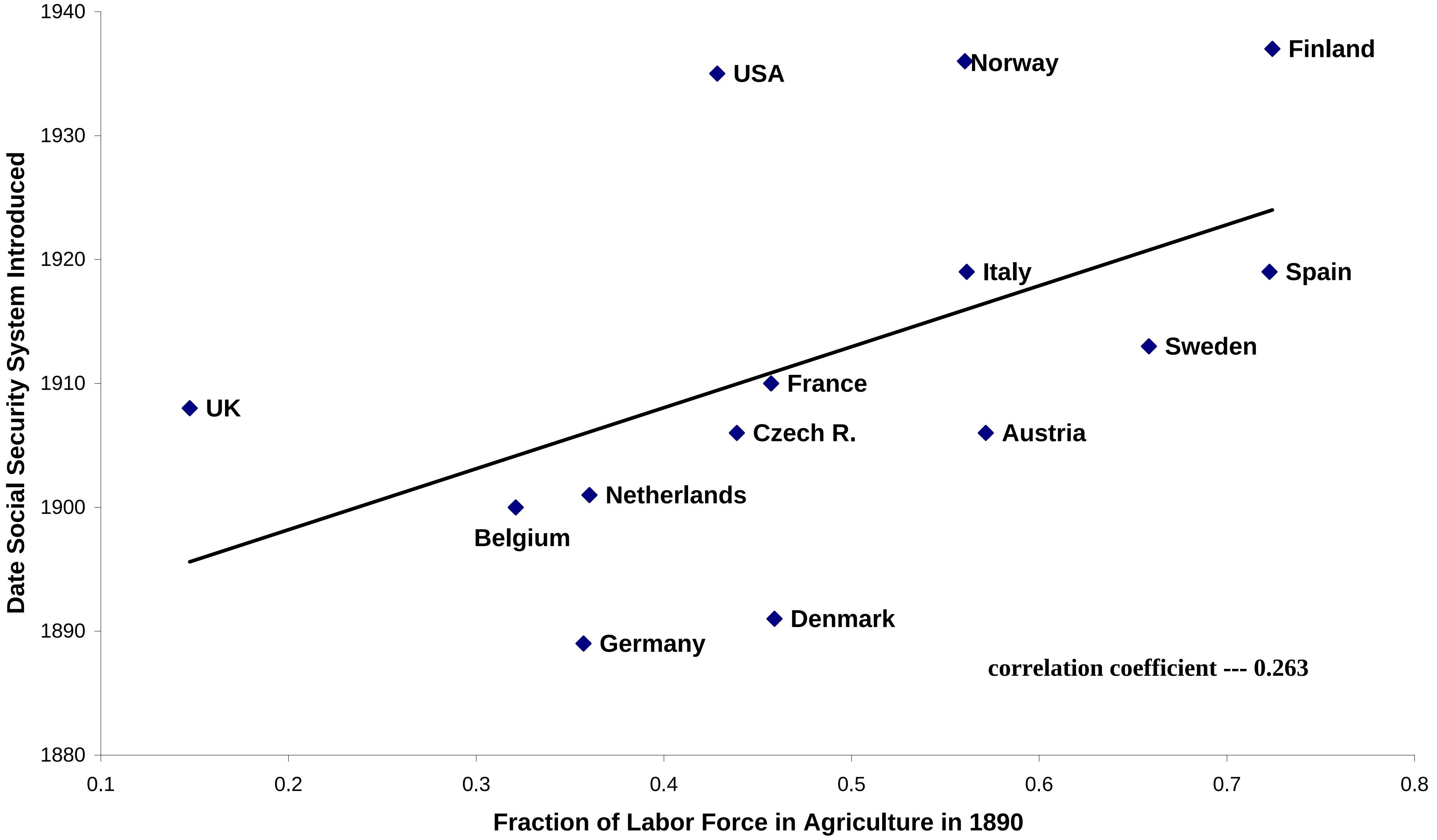


Figure 4 --- TFP in Agriculture and Non Agriculture

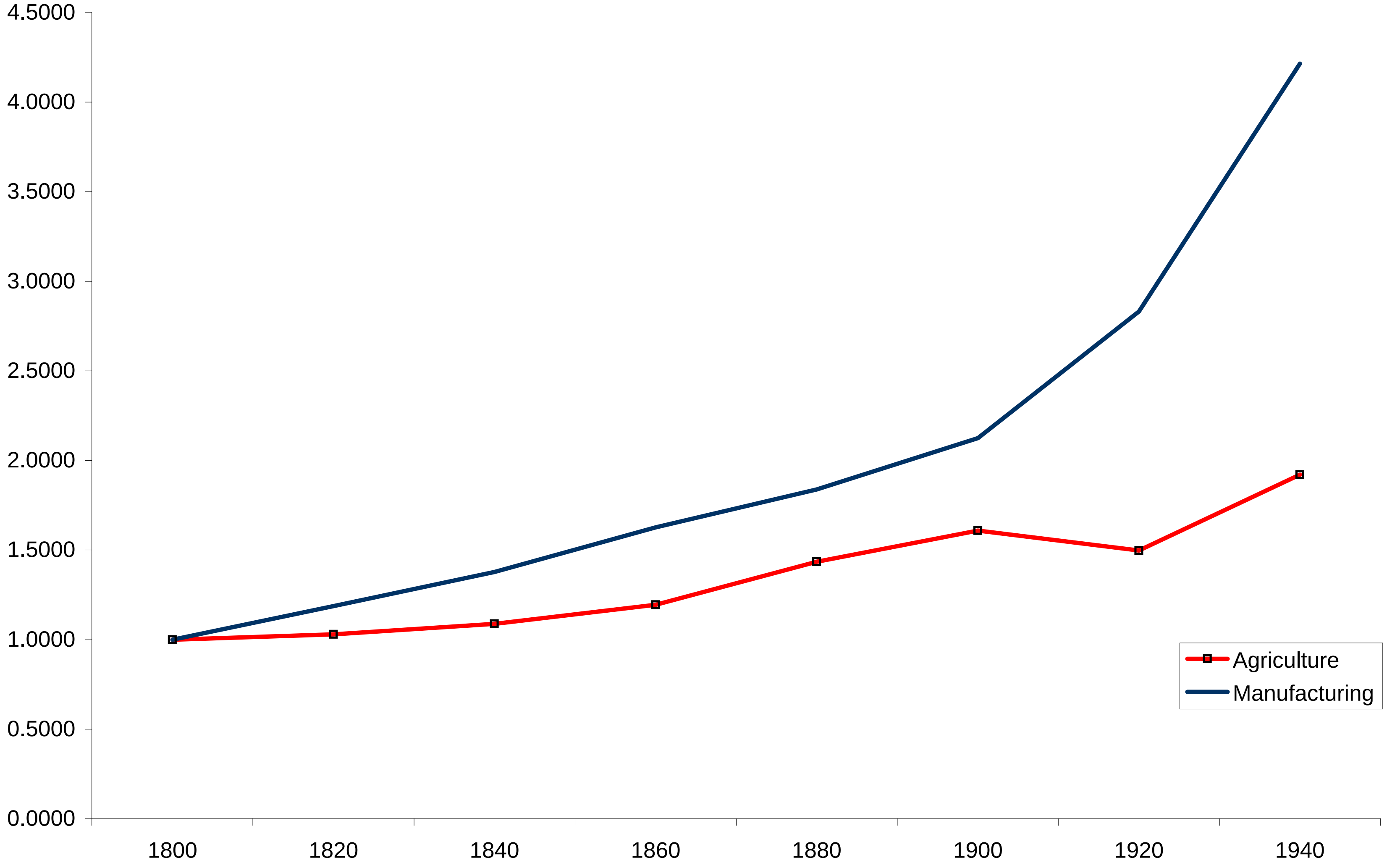
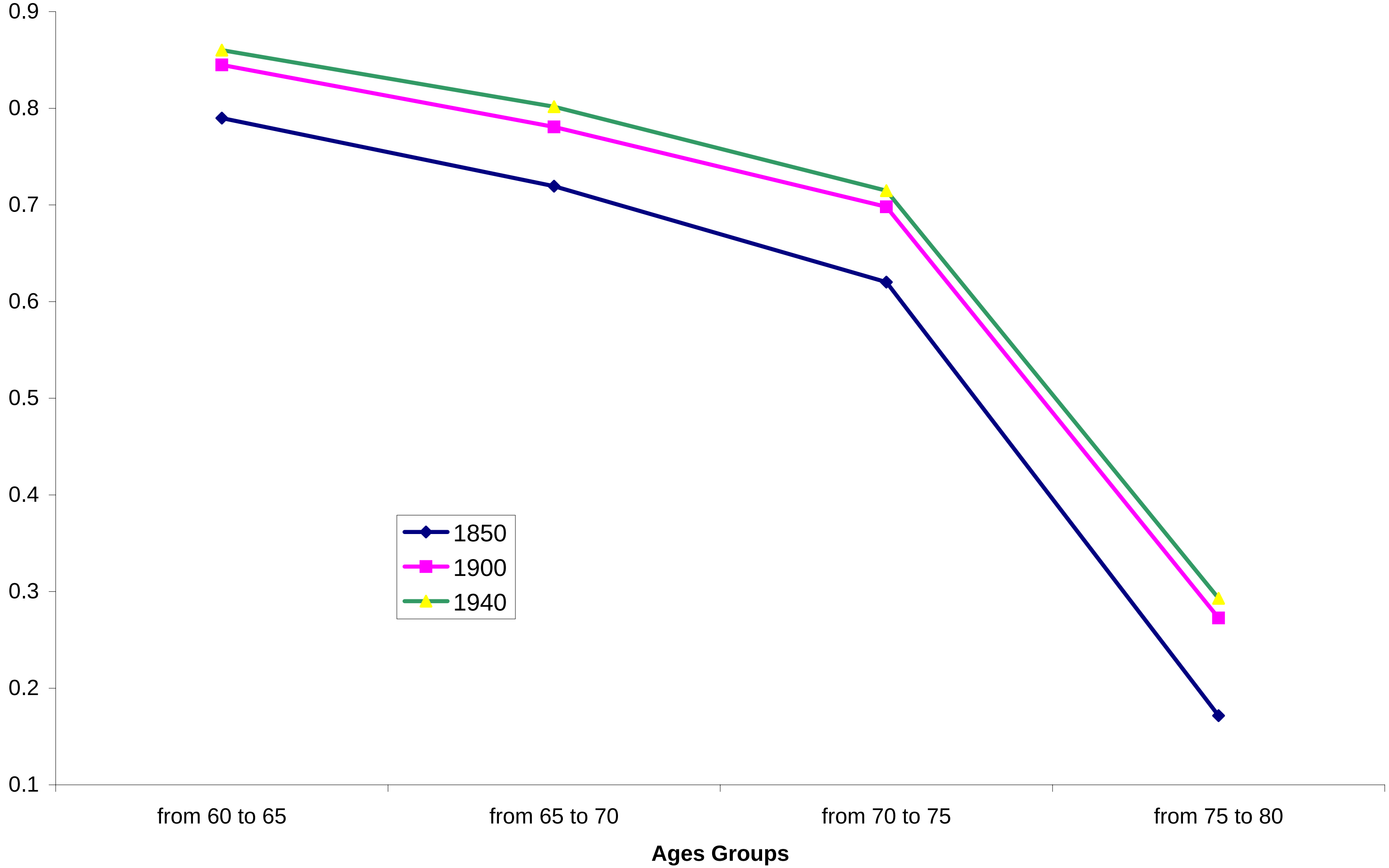


Figure 5 --- Conditional Survival Probabilities



ONLINE APPENDIX FOR

The Farm, the City, and the Emergence of Social Security

Elizabeth M. Caucutt, Thomas F. Cooley, and Nezih Guner

Appendix B - Updating and aggregating

The evolution of the asset distribution

Here we describe how G is determined. This entails updating $\psi_m^c(a)$, $\psi_o^c(a)$, $\psi_m^{f1}(a)$, $\psi_m^{f0}(a)$ and $\psi_o^f(a)$ in a manner that is consistent with the savings behavior of individuals. To this end, let $Q = [0, \overline{a}]$ be the set of possible asset holdings for an individual in this economy. First, consider next period's asset distribution among the old in the city. This distribution will be determined by the savings of the current middle-aged agents in the city who survive to the next period. Then, it must be the case that for all $\widetilde{a} \in Q$,

$$\psi_o^{c'}(\widetilde{a}) = \pi \int_Q I\{a_m^c(a,S) = \widetilde{a}\} d\psi_m^c(a), \tag{16}$$

where $I(.) = 1$ if $a_m^c(a,S) = \widetilde{a}$, and 0, otherwise. Similarly, the asset distribution of the old on the farm is

$$\psi_o^{f'}(\widetilde{a}) = \pi \int_Q I\{a_m^{f0}(a,S) = \widetilde{a}\} d\psi_m^{f0}(a) + \pi \int_Q I\{a_m^{f1}(a,S) = \widetilde{a}\} d\psi_m^{f1}(a), \tag{17}$$

where, with some abuse of notation, we use I as the appropriate indicator function.

Next period's asset distribution among the middle-aged agents in the city is determined by the location and savings decisions of young agents. One complication is that not all young agents make the same savings decisions. While some of them are born in the city, others move to the city this period. Furthermore, some of those movers had landless parents and some had landed parents. The following equation lists each of these cases:

$$\begin{aligned}
\psi_m^{c'}(\widetilde{a}) = \int_Q [\pi I\{a_y^c(a_m^c(a,S),S) = \widetilde{a}\} + (1-\pi) I\{a_y^c(a_m^c(a,S),S) + a_m^c(a,S) = \widetilde{a}\}] d\psi_m^c(a) \\
+L(a_m^{f0}(a,S),S) \int_Q [\pi I_\pi^0\{a_y^{f0g}(a_m^{f0}(a,S),S) &= \widetilde{a}\} \\
+(1-\pi) I_{1-\pi}^0\{a_y^{f0g}(a_m^{f0}(a,S),S) + a_m^{f0}(a,S) &= \widetilde{a}\}] d\psi_m^{f0}(a) \\
+L(a_m^{f1}(a,S),S) \int_Q [\pi I_\pi^1\{a_y^{f1g}(a_m^{f1}(a,S),S) &= \widetilde{a}\} \\
+(1-\pi) I_{1-\pi}^1\{a_y^{f1g}(a_m^{f1}(a,S),S) + a_m^{f1}(a,S) &= \widetilde{a}\}] d\psi_m^{f1}(a).
\end{aligned} \tag{18}$$

The first line represents the total assets held by next period's middle-aged agents, who are young this period and were also born in the city. Their savings decisions are given by $a_y^c(a_m^c(a,S),S)$. If they do not receive any bequest, which happens with probability π, these are all the assets they have. There is however a $1-\pi$ chance that they receive a bequest. In this case, their total assets consist of their own savings and their parent's assets, and are given by $a_y^c(a_m^c(a,S),S)+a_m^c(a,S)$. The next two lines consider the same cases for young agents who go to the city and have landless parents, while the last two rows do the same for those who go to the city and have landed parents.

Finally, next period's asset distribution for middle-aged agents on the farm is given by the savings decisions of the young who choose to stay there. For the landless-middle-aged farmers we have,

$$\psi_m^{f0'}(\widetilde{a}) = \pi[(1-L(a_m^{f0}(a,S),S))\int_Q I_\pi^0\{a_y^{f0s}(a_m^{f0}(a,S),S)=\widetilde{a}\}d\psi_m^{f0}(a) \tag{19}$$

$$+(1-L(a_m^{f1}(a,S),S))\int_Q I_\pi^1\{a_y^{f1s}(a_m^{f1}(a,S),S)=\widetilde{a}\}d\psi_m^{f1}(a)].$$

And, for the landed-middle-aged farmers we have,

$$\psi_m^{f1'}(\widetilde{a}) = (1-\pi)[(1-L(a_m^{f0}(a,S),S))\int_Q I_{1-\pi}^0\{a_y^{f0s}(a_m^{f0}(a,S),S)+a_m^{f0}(a,S)=\widetilde{a}\}d\psi_m^{f0}(a) \tag{20}$$

$$+(1-L(a_m^{f1}(a,S),S))\int_Q I_{1-\pi}^1\{a_y^{f1s}(a_m^{f1}(a,S),S)+a_m^{f1}(a,S)=\widetilde{a}\}d\psi_m^{f1}(a)].$$

Evolution of the age-location distribution

In order to determine H, we consider how the location decisions are updated. Suppose the current location decisions of agents are given by $\Lambda=(\lambda_y,\lambda_m,\lambda_o)$. Since all young agents survive to middle age, it must be the case that $\lambda_m'=\lambda_y$. Similarly, since the survival probability, π, is identical in both locations, $\lambda_o'=\lambda_m$. The fraction of young agents who will be on the farm, however, depends on the location decisions of those agents who are born on the farm. A fraction λ_y will be born on the farm next period. Yet, according to equation (11), some of them will move to the city. Hence, for any S', the total fraction who stay, among those whose parent does not have any land, is given by $\int(1-L(a_m^{f0}(a,S'),S'))d\psi_m^{f0'}(a)$. The same expression for those whose parent has land is given by $\int(1-L(a_m^{f1}(a,S'),S'))d\psi_m^{f1'}(a)$. Putting these pieces together implies the following consistency condition for Λ'

$$\Lambda' = \Bigg(\lambda_y\Bigg[\int(1-L(a_m^{f0}(a,S'),S'))d\psi_m^{f0'}(a)$$

$$+\int(1-L(a_m^{f1}(a,S'),S'))d\psi_m^{f1'}(a)\Bigg],\lambda_y,\lambda_m\Bigg). \tag{21}$$

Aggregation of assets

In this economy, assets are owned either by old or by middle-aged agents. Hence, given $\psi_m^c(a)$ and $\psi_o^c(a)$, the current level of aggregate assets in the city, A^c, is simply

$$A^c = (1-\lambda_m)\int a d\psi_m^c(a) + (1-\lambda_o)\int a d\psi_o^c(a). \tag{22}$$

Similarly, the aggregate asset level on the farm, A^f, is

$$A^f = (1-\pi)\lambda_m \int a d\psi_m^{f1}(a) + \pi\lambda_m \int a d\psi_m^{f0}(a) + \lambda_o \int a d\psi_o^f(a). \tag{23}$$

Given the particular demographic structure we have imposed, in order to determine the aggregate assets next period, all we need to know is the asset distribution of the middle-aged agents. To see this, note that next period's aggregate assets are determined by the savings decisions of young and middle-aged agents. Since the savings decisions of the young depend on the expected bequests and these bequests are determined by the savings of the middle-aged agents, in order to find next period's aggregate asset level $A^{c'}$, $\psi_m^c(a)$ and $\psi_m^{f\kappa}(a)$ provide sufficient information. In particular, next period's aggregate asset level in the city is given by

$$\begin{aligned} A^{c'} \quad = \quad & (1-\lambda_m)\int \left[a_y^c(a_m^c(a,S),S) + a_m^c(a,S)\right] d\psi_m^c(a) \\ & +\lambda_m[\int L(a_m^{f0}(a,S),S)a_y^{f0g}(a_m^{f0}(a,S),S)d\psi_m^{f0}(a) \\ & +\int L(a_m^{f1}(a,S),S))a_y^{f1g}(a_m^{f1}(a,S),S)d\psi_m^{f1}(a)]. \end{aligned} \tag{24}$$

The first line in this equation is the portion of next period's assets that is determined by the savings decisions of the agents in the city. Here $\int a_m^c(a,S)d\psi_m^c(a)$ gives the total savings of the middle-aged agents. These savings are either carried to their old age, or left as accidental bequests and constitute part of the assets owned by middle-aged agents next period. The term $\int a_y^c(a_m^c(a,S),S)d\psi_m^c(a)$ is the other part of the assets owned by middle-aged agents next period. It captures the savings done by the young, who in equilibrium anticipate correctly that they will receive $a_m^c(a,S)$ as bequests. The next two lines capture the part of aggregate assets in the city that come from young agents who just moved to the city. The savings decisions of these newcomers depend on their parent's asset and land holding status, and are different from those of the young agents who are born in the city. Hence, if a young farmer whose parent has a units of assets and no land decides to go to the city, then $L(a_m^{f0}(a,S),S) = 1$ and he saves $a_y^{f0g}(a_m^{f0}(a,S),S)$. The term $\int L(a_m^{f0}(a,S),S)a_y^{f0g}(a_m^{f0}(a,S),S)d\psi_m^{f0}(a)$ is the aggregation of such assets.

In a similar fashion, next period's aggregate asset level on the farm is also determined by the asset distribution of landed- and landless-middle-aged agents and by the location decisions of the young. It is given by

$$\begin{aligned} A^{f'} \quad = \quad & \lambda_m[\int[(1-L(a_m^{f0}(a,S),S))a_y^{f0s}(a_m^{f0}(a,S),S) + a_m^{f0}(a,S)]d\psi_m^{f0}(a) + \\ & \int[(1-L(a_m^{f1}(a,S),S))a_y^{f1s}(a_m^{f1}(a,S),S) + a_m^{f1}(a,S)]d\psi_m^{f1}(a)] \end{aligned} \tag{25}$$

Like equation (24), the terms $\int a_m^{f0}(a,S)d\psi_m^{f0}(a)$ and $\int a_m^{f1}(a,S)d\psi_m^{f1}(a)$ represent the total savings of the middle-aged-landless and -landed agents, respectively, while the terms $\int a_y^{f0s}(a_m^{f0}(a,S),S)d\psi_m^{f0}(a)$ and $\int a_y^{f1s}(a_m^{f1}(a,S),S)d\psi_m^{f1}(a)$ are the savings done by the young who choose to stay on the farm.

Appendix C - Ordering Preferred Tax Levels

In contrast to middle-aged-landless farmers, both middle-aged-landed farmers and middle-aged-city workers are likely to prefer a positive social security tax. As a result, when agents move to the city, middle-aged-city workers and landed-farmers form a coalition that makes up a majority that supports implementing a social security program. Understanding how the preferred tax levels of these individuals are ordered can help identify who the decisive voter might be.

Using Proposition 3, we can order middle-aged individual's preferred tax levels by their initial asset levels, within a location-land status category. For instance, among middle-aged-city workers, the optimal tax choice is increasing in initial asset level. But since each location-land status category implies a different age-income profile, it is less straightforward to compare preferred taxes across location-land status type. We can show that every middle-aged-landed farmer enters middle-age with at least as much capital as every middle-aged-landless farmer, which then implies that every landed farmer prefers a tax that is at least as great as every landless farmer. We begin by showing that if a middle-aged-landed farmer has an initial asset level that is greater than or equal to that of a middle-aged-landless farmer, then he prefers a higher tax level. Once this is established, we show that every middle-aged-landed farmer has an initial asset level that is at least as great as every middle-aged-landless farmer. Together, these imply that every landed-farmer prefers a tax that is at least as great as every landless-farmer.

Lemma 1. *Suppose that a middle-aged-landed farmer has an initial asset level of $\hat{a}$ and prefers a tax level of $\hat{\tau}$, and a middle-aged-landless farmer has an initial asset level of $\tilde{a}$ and prefers a tax level of $\tilde{\tau}$. If $\hat{a} \geq \tilde{a}$, then $\hat{\tau} \geq \tilde{\tau}$.*

Proof of Lemma 1:

First note that if $\frac{2}{\pi} < 1 + r$, then all middle-aged agents optimally choose a tax level of zero.

What if $\frac{2}{\pi} > 1 + r$? Let $\hat{\tau}$ be the preferred tax level of a landed-middle-aged farmer, who starts middle-age with assets, $\hat{a}$. Since this person is making an optimal joint decision on taxes and assets, and we know the return to taxes is higher than the return to assets, his middle-aged optimal asset choice will be zero. Since $\hat{\tau}$ is optimal for the middle-aged-landed farmer, it solves:

$$-u'\left(w^f \varepsilon_m^f + \frac{q}{\lambda} + (1+r)\hat{a} - \hat{\tau}\right) + 2\beta u'\left(w^f \varepsilon_o^f + \frac{q}{\lambda} + \frac{2\hat{\tau}}{\pi}\right) \leq 0.$$

Suppose a middle-aged-landless farmer, with assets $\tilde{a}$, faces this tax level of $\hat{\tau}$ (it may not be his optimal choice). This means that his middle-age asset choice may be positive. Let this choice be given by $a'(\hat{\tau}) \geq 0$. We can write the middle-aged-landless farmer's tax first order condition evaluated at $\hat{\tau}$:

$$-u'\left(w^f \varepsilon_m^f + (1+r)\tilde{a} - \hat{\tau} - a'(\hat{\tau})\right) + 2\beta u'\left(w^f \varepsilon_o^f + \frac{q}{\lambda} + (1+r)a'(\hat{\tau}) + \frac{2\hat{\tau}}{\pi}\right).$$

Because the middle-aged-landless farmer has strictly lower consumption in middle age than the middle-aged-landed farmer, and weakly greater consumption in old-age, his first order condition evaluated at $\hat{\tau}$ will be strictly lower than the landed's first order condition evaluated at $\hat{\tau}$. This implies that the landless-middle-aged farmer prefers a weakly lower tax level:

$$-u'\left(w^f \varepsilon_m^f + (1+r)\tilde{a} - \hat{\tau} - a'(\hat{\tau})\right) + 2\beta u'\left(w^f \varepsilon_o^f + \frac{q}{\lambda} + (1+r)a'(\hat{\tau}) + \frac{2\hat{\tau}}{\pi}\right) <$$

$$-u'\left(w^f \varepsilon_m^f + \frac{q}{\lambda} + (1+r)\hat{a} - \hat{\tau}\right) + 2\beta u'\left(w^f \varepsilon_o^f + \frac{q}{\lambda} + \frac{2\hat{\tau}}{\pi}\right) \leq 0.$$

$\square$

The landed-middle-aged farmer is only landed if his parent dies. Since his parent dies, he also receives a capital bequest (if his parent had a positive asset level). On the other hand, the landless farmer receives no bequest since his parent survived. Because the capital with which a farmer enters middle-age is the sum of his own savings while young and any accidental bequest, it turns out that the landed farmer has a higher initial asset level than the landless farmer. The highest initial asset level among landless farmers is less than or equal to the lowest initial asset level (including bequests) of the landed farmers.

Lemma 2. *Let $\bar{a}$ be the highest initial asset level of middle-age-landless farmers, and let $\underline{a}$ be the lowest initial asset level of middle-age-landed farmers, then $\bar{a} \leq \underline{a}$.*

Proof of Lemma 2

Let $c_y = I_y - \tau - a_1$, $c_m^0 = I_m^0 + (1+r)a_1 - \tau - a_2^0$, $c_m^1 = I_m^1 + (1+r)(a_1 + b) - \tau - a_2^1$, $c_o^0 = I_o + (1+r)a_2^0 + \frac{2\tau}{\pi}$, $c_o^1 = I_o + (1+r)a_2^1 + \frac{2\tau}{\pi}$, and $a_1(b)$ be the asset choice of a young farmer conditional on his parent's own middle-age saving b. Suppose:

$$-1 \leq \frac{\partial a_1}{\partial b} \leq 0.$$

This implies that a_1 (asset level with which landless-farmer enters middle-age) is decreasing in b, and $a_1 + b$ (asset level with which landed-farmer enters middle-age)is increasing in b.

Suppose young farmers have parents who save $b \in [b_{min}, b_{max}]$. Then, landless-farmers enter middle-age with $a_1 \in [a_1(b_{max}), a_1(b_{min})]$, and landed-farmers enter middle-age with $a_1 \in [a_1(b_{min}) + b_{min}, a_1(b_{max}) + b_{max}]$. Let $a_1(b_{min}) = \bar{a}$, and $a_1(b_{min}) + b_{min} = \underline{a}$. Therefore, $\bar{a} \leq \underline{a}$.

To show $-1 \leq \frac{\partial a_1}{\partial b} \leq 0$, implicitly differentiate the young farmer's first order condition in assets (Equation 32) with respect to the bequest level. Note, this expression contains the implicit derivatives of his middle-age asset choices in each state of the world (gets bequest, and does not get bequest) with respect to the bequest level.

$$\frac{\partial a_1}{\partial b} = \frac{-\beta\pi(1+r)^4 u''(c_1^o)u''(c_1^m)\left[(1-\pi)u''(c_0^m) + \beta\pi(1+r)^2 u''(c_0^o)\right]}{D},$$

where

$$\begin{gathered} D = \beta\pi(1+r)^4 u''(c_1^o)u''(c_1^m)\left[(1-\pi)u''(c_0^m) + \beta\pi(1+r)^2 u''(c_0^o)\right] \\ +\beta^2\pi^2(1+r)^4 u''(c_1^m)u''(c_0^o)u''(c_0^m) + \beta^3\pi^3(1+r)^6 u''(c_0^m)u''(c_0^o)u''(c_1^o) \\ +u''(c^y)u''(c_0^m)u''(c_1^m) + \beta\pi(1+r)^2 u''(c^y)u''(c_1^m)u''(c_0^o) \\ +\beta\pi(1+r)^2 u''(c^y)u''(c_o^m)u''(c_1^o) + \beta^2\pi^2(1+r)^4 u''(c^y)u''(c_0^o)u''(c_1^o). \end{gathered}$$

Since the numerator is positive and the denominator is negative, we have that $\frac{\partial a_1}{\partial b} \leq 0$. Note that the first two terms in D are the negative of the numerator. Therefore, we have $-1 \geq \frac{\partial a_1}{\partial b}$.

□

Proposition 5. *Let $\bar{\tau}$ be the highest preferred tax level of middle-age-landless farmers, and let $\underline{\tau}$ be the lowest preferred tax level of middle-age-landed farmers, then $\bar{\tau} \leq \underline{\tau}$.*

Proof of Proposition 5: This follows directly from Lemma 1 and Lemma 2.

□

We next turn to comparisons across sectors to highlight how the tax preferences of middle-aged-city workers relate to farmers. When it comes to optimal tax comparisons between farmers and city workers, we need to put restrictions on their initial asset levels. We show in Proposition 6 that if a middle-aged-city worker has an initial asset level that is equivalent to, or greater than, that of a middle-aged-landless farmer, then the middle-aged-city worker prefers a higher tax level. We are, however, unable to rank the initial asset levels of all city workers relative to all landless, or all landed farmers. These asset levels depend on the age-income profiles of those in the city relative to the farm, and the potential bequest levels in each place. It is reasonable to think that the young-city worker will save more in the first period than the young farmer, because the young farmer has some positive probability of inheriting land in the next period, receiving a big boost to his income. However, city workers also earn higher wage income than farmers (if not, then living on the farm would dominate living in the city and no one would live in the city). Higher income while young and middle-age has an ambiguous effect on saving while young.

Proposition 6. *Suppose that a middle-aged-city worker has an initial asset level of $\hat{a}$ and prefers a tax level of $\hat{\tau}$, and a middle-aged-landless farmer has an initial asset level of $\tilde{a}$ and prefers a tax level of $\tilde{\tau}$. If $\hat{a} \geq \tilde{a}$, then $\hat{\tau} \geq \tilde{\tau}$.*

Proof of Proposition 6:

First note that if $\frac{2}{\pi} < 1 + r$, then all middle-aged agents optimally choose a tax level of zero.

What if $\frac{2}{\pi} > 1 + r$? Let $\hat{\tau}$ be the preferred tax level of a middle-aged guy in the city, who starts the period with $\hat{a}$ assets. Since he is making an optimal joint decision on taxes and assets, and we know the return to taxes is higher than the return to assets, his optimal middle-age asset choice will be zero. Since $\hat{\tau}$ is optimal for him, it solves:

$$-u'\left(w^c \varepsilon_m^c + (1+r)\hat{a} - \hat{\tau}\right) + 2\beta u'\left(w^c \varepsilon_o^c + \frac{2\hat{\tau}}{\pi}\right) \leq 0.$$

Suppose a middle-aged-landless farmer with asset level, $\tilde{a}$, faces this tax level of $\hat{\tau}$ (it may not be his optimal choice). This means that his middle-age asset choice may be positive. Let this choice be given by $a'(\hat{\tau}) \geq 0$. We can write the middle-aged-landless farmer's first order condition with respect to the tax level, evaluated at $\hat{\tau}$:

$$-u'\left(w^f \varepsilon_m^f + (1+r)\tilde{a} - \hat{\tau} - a'(\hat{\tau})\right) + 2\beta u'\left(w^f \varepsilon_o^f + \frac{q}{\lambda} + (1+r)a'(\hat{\tau}) + \frac{2\hat{\tau}}{\pi}\right).$$

In equilibrium, if anyone is living in the city, it has to be the case that: $w^c \varepsilon_m^c \geq w^f \varepsilon_m^f$, and $w^f \varepsilon_o^f + \frac{q}{\lambda} > w^c \varepsilon_o^c$.

This implies that the middle-aged-landless farmer has strictly lower consumption in middle age than the middle-aged-city worker, and weakly greater consumption in old-age, his first order condition evaluated at $\hat{\tau}$ will be strictly lower than the city guy's first order condition evaluated at $\hat{\tau}$. This implies that the landless-middle-aged farmer prefers a weakly lower tax level:

$$-u'\left(w^f \varepsilon_m^f + (1+r)\tilde{a} - \hat{\tau} - a'(\hat{\tau})\right) + 2\beta u'\left(w^f \varepsilon_o^f + \frac{q}{\lambda} + (1+r)a'(\hat{\tau}) + \frac{2\hat{\tau}}{\pi}\right) <$$

$$-u'\left(w^c \varepsilon_m^c + (1+r)\hat{a} - \hat{\tau}\right) + 2\beta u'\left(w^c \varepsilon_o^c + \frac{2\hat{\tau}}{\pi}\right) \leq 0.$$

□

Both landed farmers and city workers have fairly constant age-income profiles, so this comparison is not as straightforward as was the case for city workers and middle-aged-landless farmers. Suppose that $\varepsilon_m^j = \varepsilon_o^j,\ j = f, c$.

In this case, both middle-aged-city workers and middle-aged-landed farmers have a flat age-earning profile, i.e. for both $I_m = I_o = I$. The following proposition shows how optimal taxes change with I.

Proposition 7. *Suppose that the optimal tax level for an individual with constant earnings, I, in middle and old age, and initial asset level, a, is $\tau^* > 0$. Suppose $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$, if $\beta > \frac{1}{2}$, then $\frac{\partial \tau^*}{\partial I} > 0$.*

Proof of Proposition 7:

Since the preferred tax level of the landed-middle-aged farmer is strictly positive, $\tau^* > 0$:

$$-u'\left(I + (1+r)a - \tau^*\right) + 2\beta u'\left(I + \frac{2\tau^*}{\pi}\right) = 0.$$

Implicitly differentiate this condition to get:

$$\frac{\partial \tau^*}{\partial I} = \frac{u''(I + (1+r)a - \tau^*) - 2\beta u''(I + \frac{2\tau^*}{\pi})}{u''(I + (1+r)a - \tau^*) + \frac{4\beta}{\pi} u''(I + \frac{2\tau^*}{\pi})}. \tag{26}$$

So, the preferred tax level is increasing in income if $u''(I + (1+r)a - \tau^*) < 2\beta u''(I + \frac{2\tau^*}{\pi})$, otherwise, it is decreasing in income. Let $c_m = I + (1+r)a - \tau^*$, and $c_o = I + \frac{2\tau^*}{\pi}$.

With CES utility we have $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$, $u'(c) = c^{1-\sigma}$, and $u''(c) = -\sigma c^{-\sigma-1}$.

If we substitute the functional forms into $u'(c_m) = 2\beta u'(c_o)$, (the first order condition determining the optimal tax level of the middle-aged individual) we find

$$2\beta = (\frac{c_o}{c_m})^{\sigma}.$$

Substitute this into the numerator of Equation 26. After some algebra, the sign of the numerator is the same as the sign of: $c_m - c_o$.

From the individual's optimal tax problem we have, $u'(c_m) = 2\beta u'(c_o)$. Since, $u''(\cdot) < 0$, if $\beta > .5$, we have $c_m < c_o$.

Therefore, $\frac{\partial \tau^*}{\partial I} > 0$.

□

Proposition 7 states that with standard CES utility and a discount factor greater than one half, which is the case we consider in our quantitative work, the preferred tax level rises with income level I. How do incomes of middle-aged-city workers and middle-aged-landed farmers compare? If anyone chooses to live on the farm, the middle-aged-city worker has lower earnings in both periods than the landed farmer, and as a result, Proposition 7 implies that the preferred tax level of a middle-aged-city worker is lower than the preferred tax level of a middle-aged-landed farmer who has the same level of initial assets. Together with Proposition 6, this implies that the preferred taxes of middle-aged-city workers are likely to fall between middle-aged-landless and -landed farmers, leaving their exact ordering a quantitative question.

Appendix D - Proofs

Proof of Proposition 1: The problem of a middle-aged agent who faces a tax level τ, and who enters the period with asset level a, is given by:

$$\max_{a' \geq 0} \{u\left(I_m + (1+r)a - \tau - a'\right) + \beta\pi u(I_o + (1+r)a' + \frac{2\tau}{\pi})\}. \tag{27}$$

This yields the following first order condition for a' :

$$-u'\left(I_m + (1+r)a - \tau - a'\right) + \beta\pi(1+r)u'(I_o + (1+r)a' + \frac{2\tau}{\pi}) \leq 0. \tag{28}$$

Let $c_m = I_m + (1+r)a - \tau - a'$, and $c_o = I_o + (1+r)a' + \frac{2\tau}{\pi}$.

i. Implicitly differentiate the middle-aged agent's first order condition (Equation 28) with respect to assets:

$$u''(c_m)[-1 - \frac{\partial a'}{\partial \tau}] = \beta\pi(1+r)u''(c_o)[(1+r)\frac{\partial a'}{\partial \tau} + \frac{2}{\pi}],$$

which yields:

$$\frac{\partial a'}{\partial \tau} = \frac{-u''(c_m) - 2\beta(1+r)u''(c_o)}{u''(c_m) + \beta\pi(1+r)^2 u''(c_o)}.$$

If $\frac{2}{\pi} > (1+r)$, then $\frac{\partial a'}{\partial \tau} < -1$.

ii. Implicitly differentiate the middle-aged agent's first order condition (Equation 28) with respect to assets:

$$u''(c_m)[(1+r) - \frac{\partial a'}{\partial a}] = \beta\pi(1+r)^2 u''(c_o)\frac{\partial a'}{\partial a},$$

which yields:

$$\frac{\partial a'}{\partial a} = (1+r)\left[\frac{u''(c_m)}{u''(c_m) + \beta\pi(1+r)^2 u''(c_o)}\right].$$

Therefore, $1 + r > \frac{\partial a'}{\partial a} > 0$.

iii. Implicitly differentiate the middle-aged agent's first order condition (Equation 28) with respect to assets:

$$u''(c_m)[1 - \frac{\partial a'}{\partial I_m}] = \beta\pi(1+r)^2 u''(c_o)\frac{\partial a'}{\partial I_m},$$

which yields:

$$\frac{\partial a'}{\partial I_m} = \frac{u''(c_m)}{u''(c_m) + \beta\pi(1+r)^2 u''(c_o)}.$$

Therefore, $1 > \frac{\partial a'}{\partial I_m} > 0$.

Implicitly differentiate the middle-aged agent's first order condition (Equation 28) with respect to assets:

$$-u''(c_m)\frac{\partial a'}{\partial I_o} = \beta\pi(1+r)u''(c_o)[(1+r)\frac{\partial a'}{\partial I_o} + 1],$$

which yields:

$$\frac{\partial a'}{\partial I_o} = \frac{-\beta\pi(1+r)u''(c_o)}{u''(c_m) + \beta\pi(1+r)^2 u''(c_o)}.$$

Therefore, $0 > \frac{\partial a'}{\partial I_o} > -1$.

Proof of Proposition 2:

The middle-age person's optimal tax problem is given by:

$$\max_{\tau \geq 0} \{u\left(I_m + (1+r)a - \tau - a'(\tau)\right) + \beta\pi u(I_o + (1+r)a'(\tau) + \frac{2\tau}{\pi})\}. \tag{29}$$

The first order condition for this problem, with the first order condition for assets substituted in, is:

$$-u'\left(I_m + (1+r)a - \tau - a'\right) + 2\beta u'(I_o + (1+r)a' + \frac{2\tau}{\pi}) \leq 0. \tag{30}$$

To prove the first part of this Proposition, we need to show that when the return to social security is strictly greater than the return to assets for a middle-aged agent, his optimal asset level is zero. We do this by showing that when $\frac{2}{\pi} > 1 + r$, the middle-aged agent's first order condition for assets (Equation (28)) is strictly negative. We know that the middle-aged agent's first order condition for the tax level (Equation (30)) is non-positive:

$$0 \geq -u'\left(I_m + (1+r)a - \tau - a'\right) + 2\beta u'(I_o + (1+r)a' + \frac{2\tau}{\pi}).$$

Using the fact that $2 > \pi(1+r)$, this can be rewritten as:

$$0 > -u'\left(I_m + (1+r)a - \tau - a'\right) + \pi(1+r)\beta u'(I_o + (1+r)a' + \frac{2\tau}{\pi}),$$

where the right hand side is just the middle-aged agent's first order condition for assets. Since this is strictly negative, the asset level is zero.

To prove the second part of the Proposition, we need to show that when the return to assets is strictly greater than the return to social security, the optimal tax level is zero. We do this by showing that when $\frac{2}{\pi} < 1 + r$, the middle-aged agent's first order condition for the tax level (Equation (30)) is strictly negative. We know that the middle-aged agent's first order condition for assets (Equation (28)) is non-positive:

$$0 \geq -u'\left(I_m + (1+r)a - \tau - a'\right) + \pi(1+r)\beta u'(I_o + (1+r)a' + \frac{2\tau}{\pi}).$$

Using the fact that $\pi(1+r) > 2$, this can be rewritten as:

$$0 > -u'\left(I_m + (1+r)a - \tau - a'\right) + 2\beta u'(I_o + (1+r)a' + \frac{2\tau}{\pi}),$$

where the right hand side is just the middle-aged agent's first order condition for the tax level. Since this is strictly negative, the tax level is zero.

Proof of Proposition 3: Let $c_m = I_m + (1+r)a - \tau - a'$, and $c_o = I_o + (1+r)a' + \frac{2\tau}{\pi}$.

i. Implicitly differentiate the middle-aged agent's first order condition (Equation 30) with respect to the tax level:

$$u''(c_m)[(1+r) - \frac{\partial \tau}{\partial a}] = 2\beta u''(c_o)\frac{\partial \tau}{\partial a}\frac{2}{\pi},$$

which yields:

$$\frac{\partial \tau}{\partial a} = \frac{u''(c_m)(1+r)}{u''(c_m) + \frac{4\beta}{\pi}u''(c_o)}.$$

Therefore, $1 + r > \frac{\partial \tau}{\partial a} > 0$.

ii. Implicitly differentiate the middle-aged agent's first order condition (Equation 30) with respect to the tax level:

$$u''(c_m)[1 - \frac{\partial \tau}{\partial I_m}] = 2\beta u''(c_o)\frac{2}{\pi}\frac{\partial \tau}{\partial I_m},$$

which yields:

$$\frac{\partial \tau}{\partial I_m} = \frac{u''(c_m)}{u''(c_m) + \frac{4\beta}{\pi}u''(c_o)}.$$

Therefore, $1 > \frac{\partial \tau}{\partial I_m} > 0$.

Implicitly differentiate the middle-aged agent's first order condition (Equation 30) with respect to assets:

$$-u''(c_m)\frac{\partial \tau}{\partial I_o} = 2\beta u''(c_o)[\frac{2}{\pi}\frac{\partial \tau}{\partial I_o} + 1],$$

which yields:

$$\frac{\partial \tau}{\partial I_o} = \frac{-2\beta u''(c_o)}{u''(c_m) + \frac{4\beta}{\pi}u''(c_o)}.$$

Therefore, $0 > \frac{\partial \tau}{\partial I_o} > -1$.

Proof of Proposition 4:

Let the earnings of the young person be given by I_y. Let I_m^0 be middle-aged earnings when there is no bequest, and let I_m^1 be middle-aged earnings when there is a bequest. Since city workers only earn wage income in each state, they have $I_m^0 = I_m^1$, while I_m^0 and I_m^1 are different for young farmers since in the latter case they receive extra income from land.

The young person who faces a tax level τ, and has a possible bequest of $b(\tau)$, solves the following problem:

$$\begin{aligned}\max_{a_1 \geq 0,\ a_2^0 \geq 0,\ a_2^1 \geq 0} &\{u\left(I_y - \tau - a_1\right) + \beta\pi u\left(I_m^0 + (1+r)a_1 - \tau - a_2^0\right) \\ &+\beta(1-\pi)u\left(I_m^1 + (1+r)(a_1 + b(\tau)) - \tau - a_2^1\right) \\ &+\beta^2\pi^2 u(I_o + (1+r)a_2^0 + \frac{2\tau}{\pi}) \\ &+\beta^2\pi(1-\pi)u(I_o + (1+r)a_2^1 + \frac{2\tau}{\pi})\}.\end{aligned} \tag{31}$$

This problem yields the following first order conditions in a_1, a_2^0, and a_2^1, respectively

$$-u'\left(I_y - \tau - a_1\right) + \beta\pi(1+r)u'\left(I_m^0 + (1+r)a_1 - \tau - a_2^0\right) + \tag{32}$$

$$\beta(1-\pi)(1+r)u'\left(I_m^1 + (1+r)(a_1 + b(\tau)) - \tau - a_2^1\right) \leq 0,$$

$$-u'\left(I_m^0 + (1+r)a_1 - \tau - a_2^0\right) + \beta\pi(1+r)u'(I_o + (1+r)a_2^0 + \frac{2\tau}{\pi}) \leq 0, \tag{33}$$

and

$$-u'\left(I_m^1 + (1+r)(a_1 + b(\tau)) - \tau - a_2^1\right) + \beta\pi(1+r)u'(I_o + (1+r)a_2^1 + \frac{2\tau}{\pi}) \leq 0, \tag{34}$$

If the young agent is choosing his most preferred tax level, the problem he solves is:

$$\begin{aligned}
&\max_{\tau\geq 0}\{u\left(I_y-\tau-a_1(\tau)\right)+\beta\pi u\left(I_m^0+(1+r)a_1(\tau)-\tau-a_2^0(\tau)\right)\\
&+\beta(1-\pi)u\left(I_m^1+(1+r)(a_1(\tau)+b(\tau))-\tau-a_2^1(\tau)\right)\\
&+\beta^2\pi^2u(I_o+(1+r)a_2^0(\tau)+\frac{2\tau}{\pi})\\
&+\beta^2\pi(1-\pi)u(I_o+(1+r)a_2^1(\tau)+\frac{2\tau}{\pi})\}.
\end{aligned} \tag{35}$$

The first order condition for this problem, with the first order conditions for assets substituted in, is:

$$\begin{aligned}
-u'\left(I_y-\tau-a_1\right)-\beta\pi u'\left(I_m^0+(1+r)a_1-\tau-a_2^0\right)-\beta(1-\pi)u'\left(I_m^1+(1+r)(a_1+b(\tau))-\tau-a_2^1\right)+\\
2\beta^2\pi u'(I_o+(1+r)a_2^0+\frac{2\tau}{\pi})+2\beta^2(1-\pi)u'(I_o+(1+r)a_2^1+\frac{2\tau}{\pi})\\
+\beta(1-\pi)(1+r)\frac{\partial b}{\partial\tau}u'\left(I_m^1+(1+r)(a_1+b(\tau))-\tau-a_2^1\right)\leq 0.
\end{aligned} \tag{36}$$

We begin by proving that all three asset levels from a young person's problem, $a_1,\ a_2^0$, and a_2^1, cannot all be strictly positive if the return to social security is greater than the return to assets for a young agent. We do this in three parts.

First, show that if $a_1>0$ and $a_2^0>0$, then $a_2^1=0$. Since a_1 and a_2^0 are strictly positive, we know Equation (32) and Equation (33) both hold with equality. Solve Equation (32) for $u'\left(I_y-\tau-a_1\right)$, and Equation (33) for $u'(I_o+(1+r)a_2^0+\frac{2\tau}{\pi})$, and substitute these into the first order condition for τ, Equation (36), to get:

$$\begin{aligned}
u'\left(I_m+(1+r)a_1-\tau-a_2^0\right)\left(\frac{\pi}{(1+r)(1-\pi)}\right)\left(\frac{2}{\pi}-(1+r)^2-(1+r)\right)\\
-u'\left(I_m+(1+r)(a_1+b(\tau))-\tau-a_2^1\right)\left(1+(1+r)\left(1-\frac{\partial b}{\partial\tau}\right)\right)+2\beta u'\left(I_o+(1+r)a_2^1+\frac{2\tau}{\pi}\right)\leq 0.
\end{aligned}$$

Because the return to social security is strictly greater than the return to assets, the first piece of the above inequality is strictly positive. This implies that:

$$-u'\left(I_m+(1+r)(a_1+b(\tau))-\tau-a_2^1\right)\left(1+(1+r)\left(1-\frac{\partial b}{\partial\tau}\right)\right)+2\beta u'\left(I_o+(1+r)a_2^1+\frac{2\tau}{\pi}\right)<0.$$

The expression, $\left(1+(1+r)\left(1-\frac{\partial b}{\partial\tau}\right)\right)$, is strictly positive. Therefore:

$$-u'\left(I_m+(1+r)(a_1+b(\tau))-\tau-a_2^1\right)+\frac{2\beta}{\left(1+(1+r)\left(1-\frac{\partial b}{\partial\tau}\right)\right)}u'\left(I_o+(1+r)a_2^1+\frac{2\tau}{\pi}\right)<0.$$

Since $\frac{2}{\pi}+(1+r)^2\frac{\partial b}{\partial\tau}>1+r+(1+r)^2$, we have $\frac{2}{\left(1+(1+r)\left(1-\frac{\partial b}{\partial\tau}\right)\right)}>\pi(1+r)$. This implies that:

$$-u'\left(I_m+(1+r)(a_1+b(\tau))-\tau-a_2^1\right)+\beta\pi(1+r)u'\left(I_o+(1+r)a_2^1+\frac{2\tau}{\pi}\right)<0,$$

which means the first order condition for a_2^1 holds with strict inequality, so that $a_2^1 = 0$.

Second, show that if $a_1 > 0$ and $a_2^1 > 0$, then $a_2^0 = 0$. Since a_1 and a_2^1 are strictly positive, we know Equation (32) and Equation (34) both hold with equality. Solve Equation (32) for $u'\left(I_y - \tau - a_1\right)$, and Equation (34) for $u'(I_o + (1+r)a_2^1 + \frac{2\tau}{\pi})$, and substitute these into the first order condition for τ, Equation (36), to get:

$$u'\left(I_m + (1+r)(a_1 + b(\tau)) - \tau - a_2^1\right)\left(\frac{1-\pi}{1+r}\right)\left(\frac{2}{\pi} + (1+r)^2\frac{\partial b}{\partial \tau} - (1+r) - (1+r)^2\right)$$

$$-u'\left(I_m + (1+r)a_1 - \tau - a_2^0\right)\pi(1+(1+r)) + 2\beta\pi u'\left(I_o + (1+r)a_2^0 + \frac{2\tau}{\pi}\right) \leq 0.$$

Because the return to social security is strictly greater than the return to assets, the first piece of the above inequality is strictly positive. This implies that:

$$-u'\left(I_m + (1+r)a_1 - \tau - a_2^0\right)\pi(1+(1+r)) + 2\beta\pi u'\left(I_o + (1+r)a_2^0 + \frac{2\tau}{\pi}\right) < 0.$$

Dividing by, $\pi(1+(1+r))$, yields:

$$-u'\left(I_m + (1+r)a_1 - \tau - a_2^0\right) + \frac{2\beta}{(1+(1+r))}u'\left(I_o + (1+r)a_2^0 + \frac{2\tau}{\pi}\right) < 0.$$

Since $\frac{2}{\pi} + (1+r)^2\frac{\partial b}{\partial \tau} > 1 + r + (1+r)^2$, we have $\frac{2}{\pi} > 1 + r + (1+r)^2$, or $\frac{2}{1+(1+r)} > \pi(1+r)$. This implies that:

$$-u'\left(I_m + (1+r)a_1 - \tau - a_2^0\right) + \beta\pi(1+r)u'\left(I_o + (1+r)a_2^0 + \frac{2\tau}{\pi}\right) < 0,$$

which means the first order condition for a_2^0 holds with strict inequality, so that $a_2^0 = 0$.

Third, show that if $a_2^0 > 0$ and $a_2^1 > 0$, then $a_1 = 0$. Since a_2^0 and a_2^1 are strictly positive, we know Equation (33) and Equation (34) both hold with equality. Solve Equation (33) for $u'\left(I_o + (1+r)a_2^0 + \frac{2\tau}{\pi}\right)$, and Equation (34) for $u'(I_o + (1+r)a_2^1 + \frac{2\tau}{\pi})$, and substitute these into the first order condition for τ, Equation (36), to get:

$$-u'\left(I_y - \tau - a_1\right) + \beta\pi(1+r)u'\left(I_m + (1+r)a_1 - \tau - a_2^0\right)\left(\frac{2}{\pi(1+r)^2} - \frac{1}{1+r}\right)$$

$$+\beta(1-\pi)(1+r)u'\left(I_m + (1+r)(a_1 + b(\tau)) - \tau - a_2^1\right)\left(\frac{2}{\pi(1+r)^2} - \frac{1}{1+r} + \frac{\partial b}{\partial \tau}\right) \leq 0.$$

Because $\frac{2}{\pi} + (1+r)^2\frac{\partial b}{\partial \tau} > (1+r) + (1+r)^2$, it follows that $\frac{2}{\pi(1+r)^2} - \frac{1}{1+r} + \frac{\partial b}{\partial \tau} > 1$, and $\frac{2}{\pi(1+r)^2} - \frac{1}{1+r} > 1$. Therefore, we have the following:

$$-u'\left(I_y - \tau - a_1\right) + \beta\pi(1+r)u'\left(I_m + (1+r)a_1 - \tau - a_2^0\right)$$

$$+\beta(1-\pi)(1+r)u'\left(I_m + (1+r)(a_1 + b(\tau)) - \tau - a_2^1\right) < 0,$$

which means the first order condition for a_1 holds with strict inequality, and $a_1 = 0$.

The second piece of the proof entails showing that when the return to social security is strictly less than the return to assets, the tax optimal tax level is zero.

Solve the first order condition for a_1 (Equation (32)) for $u'\left(I_y-\tau-a_1\right)$, the first order condition for a_2^0 (Equation (33)) for $u'\left(I_o+(1+r)a_2^0+\frac{2\tau}{\pi}\right)$, and the first order condition for a_2^1 (Equation (34)) for $u'\left(I_o+(1+r)a_2^1+\frac{2\tau}{\pi}\right)$, and substitute these weak inequalities into the first order condition for the tax level (Equation (36)), to get:

$$
\begin{aligned}
-u'\left(I_y-\tau-a_1\right)-\beta\pi u'\left(I_m+(1+r)a_1-\tau-a_2^0\right)-\beta(1-\pi)u'\left(I_m+(1+r)(a_1+b(\tau))-\tau-a_2^1\right)+\\
2\beta^2\pi u'(I_o+(1+r)a_2^0+\frac{2\tau}{\pi})+2\beta^2(1-\pi)u'(I_o+(1+r)a_2^1+\frac{2\tau}{\pi})\\
+\beta(1-\pi)(1+r)\frac{\partial b}{\partial\tau}u'\left(I_m+(1+r)(a_1+b(\tau))-\tau-a_2^1\right)\\
\leq\frac{\beta\pi}{1+r}u'\left(I_m+(1+r)a_1-\tau-a_2^0\right)\left(\frac{2}{\pi}-(1+r)^2-(1+r)\right)+\\
\frac{\beta(1-\pi)}{1+r}u'\left(I_m+(1+r)(a_1+b(\tau))-\tau-a_2^1\right)\left(\frac{2}{\pi}+(1+r)^2\frac{\partial b}{\partial\tau}-(1+r)^2-(1+r)\right).
\end{aligned}
$$

The fact that $\frac{2}{\pi}<(1+r)+(1+r)^2$, implies that $\frac{2}{\pi}+(1+r)^2\frac{\partial b}{\partial\tau}<(1+r)+(1+r)^2$, which means the right hand side of the above expression is strictly negative. Therefore, the left hand side of the above expression, or the first order condition with respect to τ is negative, and $\tau=0$.

Appendix E - Can Social Security Taxes be Negative?

We have imposed a non-negativity constraint on the tax level. If we consider the optimal tax problem of the middle-aged individual, but allow for the tax level to be negative we have the following first order condition

$$-u'\left(I_m+(1+r)a-\tau-a'\right)+2\beta u'(I_o+(1+r)a'+\frac{2\tau}{\pi})=0.$$

If $\frac{2}{\pi}>1+r$, then according to Proposition 2, $a'=0$, leaving[37]

$$-u'\left(I_m+(1+r)a-\tau\right)+2\beta u'(I_o+\frac{2\tau}{\pi})=0.$$

In order for the agent to optimally choose a non-negative tax level, this first order condition must be strictly positive when $\tau=0$, i.e. the following must hold:

$$-u'\left(I_m+(1+r)a\right)+2\beta u'(I_o)>0.$$

Clearly, what is important to ensure the middle-aged agent chooses a non-negative tax is that he consumes relatively less when he is old, and so would like to move resources to his old age. This is more likely to happen if

[37]If the return to capital is greater than social security, middle-age individuals will choose an infinite negative tax, save it all, and get a higher return from that saving than they have to pay back in social security tax the next period.

β is relatively large (people value the future) and the age-income profile is relatively flat (people are more likely to want to save).

If we assume a constant elasticity of substitution utility (CES), $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$, this inequality reduces to

$$(2\beta)^{-\frac{1}{\sigma}} < \frac{I_m + (1+r)a}{I_o}.$$

If $\beta = \frac{1}{2}$, then income in middle-age needs to be at least as great as income in old age in order for the agent to choose a non-negative tax level. Not surprisingly, as the discount factor (β) rises, the ratio between middle-age income and old-age income needed to ensure a non-negative tax falls. But, the more the individual desires to smooth consumption across periods (σ), the higher this income ratio needs to be for the agent to choose a positive tax level.

What determines the $\frac{I_m}{I^0}$ ratio? For city workers, who earn $w^c \varepsilon_m^c$ when they are middle-aged and $w^c \varepsilon_o^c$ when they are old, and for landed farmers, who earn $w^f \varepsilon_m^f + \frac{q}{\lambda}$ and $w^f \varepsilon_o^f + \frac{q}{\lambda}$, this ratio depends on their life-cycle efficiency units. The larger their middle aged labor income relative to their old age labor income, the more likely they are to prefer a positive tax level. For a landless farmer, however, on top of their life-cycle efficiency units, this ratio depends crucially on the size of the return to land since his middle and old age incomes are $w^f \varepsilon_m^f$ and $w^f \varepsilon_o^f + \frac{q}{\lambda}$, respectively. The more important land is to production, the higher its return, leading to a steeper age-income profile, and downward pressure on the optimal social security tax level.

It is important to note that negative social security, which would transfer resources to middle-aged and young from the old, is not implementable with a reputational mechanism. A middle-aged median voter will vote for a positive social security tax in order to get social security payments next period. Suppose now that the middle-aged median voter prefers a transfer from the old. The old cannot be convinced to transfer resources to the middle-age and young. They are in the last period of their life and have no altruism towards their children. There are no periods left in which to impose a punishment on them if they do not comply.

Appendix F - Political economy

In order to define the political economy equilibrium, we update the aggregate state of the economy, so that $S = (\Psi, \Lambda, \tau_{-1}, h_{-1})$, where τ_{-1} is the social security tax and h_{-1} is an indicator of whether the last period's median voter deviated (i.e. introducing or keeping social security was optimal for the median voter last period, but he did not do so). Hence if $\tau_{-1} = h_{-1} = 0$, there is no social security and it can start this period. If $\tau_{-1} > 0$ and $h_{-1} = 0$, there is a social security system and it can continue this period. Finally, if $\tau_{-1} = 0$ and $h_{-1} = 1$, the median voter did not start or chose not to continue a system that was optimal for him (i.e. he deviated), and the system cannot start next period in order to punish the last period's median voter.

We represent the evolution of the political state by the function P. The role of the function P is to determine a state-contingent social security system. In particular, we assume that the social security tax level for the current period, τ, and the indicator for the next period h, are given by $(\tau, h) = P(\Psi, \Lambda, \tau_{-1}, h_{-1})$, and agents take the policy rule P as given when making their economic decisions. Note that a system with a social security tax τ that never changes is trivially defined by $(\tau, 0) = P(S)$ for all S. Let $P_\tau(S)$ represent such a system.

Definition 1. *For any $\tau_{-1} > 0$ and $h_{-1} = 0$, we will say that a policy function $P(S)$ is sustainable in state $S = (\Psi, \Lambda, \tau_{-1}, 0)$, if*

$$V^M(\Psi, \Lambda, \tau_{-1}, 0; P) \geq \mathcal{V}^M(\Psi, \Lambda)$$

where V^M is the remaining lifetime utility of the median voter in an economy with current aggregate state $S = (\Psi, \Lambda, \tau_{-1}, 0)$ and policy function P, and $\mathcal{V}^M$ is the remaining lifetime utility of the median voter if social security is eliminated forever.

The value $\mathcal{V}^M(\Psi, \Lambda)$ only depends on Ψ and Λ, i.e. the aggregate state (the distribution of physical capital and the distribution of agents between the city and the farm) in which the social security tax is eliminated. In other words, P is sustainable in S if a majority of voters vote "yes" for keeping τ_{-1} today with tomorrow's taxes determined by P, instead of moving to an economy with no social security. Let the indicator function $M(S; P)$ denote the associated yes/no decision of the median voter, i.e.

$$M(\Psi, \Lambda, \tau_{-1}, 0; P) = \begin{cases} 1, \text{ if } V^M(\Psi, \Lambda, \tau_{-1}, 0; P) \geq \mathcal{V}^M(\Psi, \Lambda) \\ 0, \text{ otherwise} \end{cases}.$$

A median voter considering a future without social security takes into account the resulting rise in aggregate capital stock and the decline in the rate of return. The decline in the rate of return gives the median voter an additional reason (besides reputation) to keep an existing system.

Suppose today's state is $\tau_{-1} = 0$ and $h_{-1} = 1$, i.e. the social security system was optimal but was not implemented. In this case, there cannot be a social security system today, i.e.

$$P(\Psi, \Lambda, 0, 1) = (0, 0). \tag{37}$$

A median voter who deviates is punished only for one period, i.e. there is no voting next period. The following period, social security can start again.

If today's political state is $\tau_{-1} > 0$ and $h_{-1} = 0$, then there is an existing social security system. In this case, the current generation simply takes a yes/no vote and the system either continues at the same tax level or ends because of a no vote, i.e.

$$P(\Psi, \Lambda, \tau_{-1}, 0) = \begin{cases} (\tau_{-1}, 0), \text{ if } M(\Psi, \Lambda, \tau_{-1}, 1; P) = 1 \\ (0, 1), \text{ if } M(\Psi, \Lambda, \tau_{-1}, 1; P) = 0 \end{cases}. \tag{38}$$

If $M(\Psi, \Lambda, \tau_{-1}, 0; P) = 0$, then the median voter ends the system. It will remain non-operative next period, but can start again after that (see Equation 37). As a result the median voter that votes no against the existing

τ_{-1}, with the hope that a social security system can start next period, is punished. Again after this punishment, the system can restart, if it is optimal for the future median voters to do so.

When $\tau_{-1} = 0$ and $h_{-1} = 0$, a social security system is not operating. It may, or may not start today, depending on the preferences of the median voter. Let $\widehat{\tau}(S)$ be the proposal by the median voter at state S. Furthermore, let $\tau^* = \arg\max_\tau V^M(S;P)$ be the optimal tax level chosen by the median voter under the constant policy rule P. We specify P such that

$$P(\Psi, \Lambda, 0, 0) = \begin{cases} (0,0), \text{ if } \widehat{\tau} = \tau^* = \arg\max_\tau V^M(S;P) = 0 \\ (\tau^*, 0), \text{ if } \widehat{\tau} = \tau^* = \arg\max_\tau V^M(S;P) > 0 \\ (0,1), \text{ if } \widehat{\tau} \neq \tau^* \end{cases} . \tag{39}$$

The current median voter might optimally choose a positive tax, and the system starts. He might optimally choose a zero tax, and the system does not start. In this case, tomorrow's median voter may start social security. If he proposes a tax level that is not his optimal choice, this is a deviation and then the system will not be available next period.

A political equilibrium is then a recursive competitive equilibrium with the policy function P defined by equations (37), (38), and (39).

Appendix G - Transition

Table VI illustrates the transitional dynamics. Computing the transition is non-trivial. Not only do the capital stock and location choices (and hence prices) have to be consistent with individual asset accumulation and migration decisions, but the sequence of tax levels that individuals expect must be those that the median voter in each generation chooses. We assume that the economy is at its 1800 steady state initially (period 0) and suddenly and unexpectedly productivity and life expectancy increase to their 1940 values. In the period of the change (period 1), the capital stock is fixed at its initial steady state level. However, due to the higher productivity in the city and the higher survival probability, the city is a much more attractive location for young farmers and many choose to migrate, $\lambda_y = 0.18$. This population shift alters the labor supply on the farm and in the city. Indeed, since a large fraction of population migrates in the first period of the transition, both farm and city wages rise. Given the rise in productivity levels, because the capital stock is fixed at its old steady state level, the return to capital increases significantly from 2.47 to 5.68. As people start moving away from the farm, the return to land starts to fall as well.

Table VI - Transition

	0	1	2	3	4	5	6	7
τ	0	0	0.09	0.09	0.09	0.09	0.09	0.09
Farm population	1	0.670	0.368	0.184	0.184	0.184	0.184	0.184
$1+r$	2.466	5.682	3.685	3.483	3.261	3.192	3.155	3.141
w^f	0.311	0.324	0.426	0.518	0.522	0.523	0.524	0.524
w^c	.178	0.378	0.504	0.524	0.547	0.555	0.559	0.561
q	0.384	0.268	0.179	0.123	0.124	0.124	0.124	0.124
K	0.052	0.052	0.165	0.223	0.249	0.257	0.262	0.264

Because the migration only affects the location of the young, in period 1 the median voter is still a middle-aged landless farmer, who prefers no social security.[38] So, in the initial period of the change, the tax remains unchanged at 0. However, agents are aware that the mass migration of young farmers to the city will shift the identity of the median voter in the next period, and alter support for social security. In the second period of the transition, the initial young migrants now become middle-aged-city workers, who support a positive (sustainable) level of social security, $\tau = .09$. After the third period, migration stops and the fraction of young farmers remains at 0.18. However, the new steady state farm population takes three periods to attain, as the initial young migrants age. As the population reallocates between the two locations and people start accumulating capital, the return to capital falls to 3.69, and then converges to 3.14 in the new steady state.[39]

[38] We computationally verify that along the transition preferences are single peaked in each period.

[39] Note that λ_y (farm population after period 2) in Table VII is less than its value in Table IV (18% versus 23%). Associated with this, the final tax rate is lower than the one in Table IV (0.09 versus 0.10). This happens since migration overshoots along the transition and we do not allow these agents to go back to the farm. When we compute the final steady state directly, this does not happen.