



Universidad
Carlos III de Madrid



This is a postprint version of the following published document:

Erosa, A., Fuster, L. and Kambourov, G. (2016), Towards a Micro-Founded Theory of Aggregate Labour Supply. [Review of Economic Studies](#), v.83, n.3, pp. 1001-1039

Available in : <https://doi.org/10.1093/restud/rdw010>

Towards a Micro-Founded Theory of Aggregate Labour Supply

ANDRÉS EROSA

Universidad Carlos III de Madrid

LUISA FUSTER

Universidad Carlos III de Madrid

and

GUEORGUI KAMBOUROV

University of Toronto

We build a heterogeneous agents life cycle model that captures a large number of salient features of individual male labour supply over the life cycle, by education, both along the intensive and extensive margins. The model provides an aggregation theory of individual labour supply, firmly grounded on individual-level micro-evidence, and is used to study the aggregate labour supply responses to changes in the economic environment. We find that the aggregate labour supply elasticity to a transitory wage shock is 1.75, with the extensive margin accounting for 62% of the response. Furthermore, we find that the aggregate labour supply elasticity to a permanent-compensated wage change is 0.44.

Key words: Aggregate labour supply, intensive margin, extensive margin, life cycle, fixed cost of work, non-linear earnings, precautionary savings, preference heterogeneity

JEL Codes: D9, E2, E13, E62, J22

1. INTRODUCTION

How responsive is aggregate labour supply to changes in the economic environment? While the answer to this question is crucial for understanding the effects of government policies and business cycles, it has led to substantial debate in the economics profession. At the core of this debate is the inconsistency between the small labour supply elasticities for men estimated in micro-studies and the large ones used in macro-models. The empirical evidence indicates that aggregate labour supply responses are determined by individual responses along both the intensive and the extensive margins.¹ While the consensus in the micro-literature is that the extensive margin, at

1. See Cooley (1995) for evidence on the adjustment in labour supply along both margins over the business cycle and Blundell *et al.* (2011) for recent evidence on the importance of both margins over time in the U.S., the U.K., and France.

the annual level, is not important for understanding labour supply responses of males, the key idea motivating our article is that individuals frequently use the extensive margin to adjust their labour supply within a year. Although once aggregated to the annual level these labour supply responses are not prominent, we argue that fluctuations in employment within a year are important for understanding labour supply responses.

Furthermore, our article builds on the idea that a rich theory of heterogeneity is needed for studying how the extensive margin matters for aggregate labour supply responses. Economic theory implies that labour supply responses along the intensive and extensive margins are distinct objects. The intensive margin responses are mainly driven by the intertemporal substitution of labour (the Frisch elasticity of labour supply). The extensive margin responses, on the other hand, are unrelated to the preference parameter typically estimated in micro-studies. Chang and Kim (2006) build on the insights from the model of indivisible labour in Hansen (1985) and Rogerson (1988), introduce heterogeneity, and show that the slope of the aggregate labour supply schedule is determined by the distribution of reservation wages rather than by the willingness to substitute leisure intertemporally, establishing that when the extensive margin is operative heterogeneity and aggregation play a crucial role in determining aggregate labour supply responses. Therefore, we build a life-cycle model with heterogeneous agents making labour supply decisions, both along the intensive and extensive margins, at subannual periods. This rich in heterogeneity model provides an aggregation theory of individual labour supply that is firmly grounded on individual-level micro-evidence and is used to study aggregate labour supply responses and the importance of the responses along the extensive margin.

We discipline our theory with a large number of facts on labour supply at the individual level documented on data from the Survey of Income and Program Participation (SIPP) and the Panel Study of Income Dynamics (PSID). Most importantly, the SIPP data reveal a quite operative extensive margin at high frequencies, such as four-month periods referred to as *quadrimesters*: the probability of entering non-employment is high, most non-employment spells are short-lived, and this pattern is observed at all stages of the life cycle. Thus, individuals frequently use the extensive margin to adjust their labour supply within a year, even at old ages. Capturing this pattern is central to the analysis in this article, and we find that it is essential for understanding labour supply responses. In addition, the PSID data reveal another insightful pattern: there are large differences in lifetime labour supply across individuals, even conditional on education and permanent income, implying that preference heterogeneity is an important factor for understanding the variation in labour supply across individuals.

We incorporate into the theoretical framework several key features that allow for an operative extensive margin of labour supply. The evidence indicates that workers rarely choose to work small number of hours, an observation that has motivated Cogan (2001) to model fixed costs of work. Moreover, there is also evidence that earnings are a convex (non-linear) function of hours of work which also provides a rationale for workers not to choose short hours (Moffitt, 1984; Gustman and Steinmeier, 1985, 1986; Keane and Wolpin, 2001; Aaronson and French, 2004). Our model incorporates both fixed costs of work and non-linear earnings, which are arguably the two main approaches in the literature to model the extensive margin. By combining these two approaches, we are able to show that there is a complementarity between fixed costs of work and non-linear earnings that enhances extensive-margin labour supply responses. The calibration implies that individuals face substantial fixed costs of work, non-linear earnings, a job finding rate of less than one when non-employed and that there is significant preference heterogeneity in the taste for leisure. All of these features are needed for the model to be consistent with the following moments: (1) the age-profile of employment rates, (2) the probability of transitioning from employment to non-employment at various stages in the life cycle, (3) the duration distribution of non-employment spells, and (4) the lifetime inequality in labour supply. Thus, our calibration

strategy ensures that the theory is consistent with the incidence and volatility of non-employment spells and the lifetime dispersion in labour supply—an important step in the development of a quantitative theory of labour supply responses.

We find that the aggregate labour supply elasticity to a one period unanticipated small wage change in the baseline economy is 1.75, with an extensive margin elasticity of 1.08 (about 62% of the aggregate labour supply response) and an intensive margin elasticity of 0.67. The empirical literature on labour supply has provided a broad set of estimates for the intertemporal elasticity of labour supply along the intensive margin. The implications of our theory are consistent with the findings in this literature.² Furthermore, we stress that the intertemporal elasticity in our model economy is computed from an ideal experiment for which there is no clear counterpart in the data: we simulate a one period small (purely) unanticipated wage change and compute aggregate labour supply responses to measure the intertemporal elasticity of labour supply. When we measure the intertemporal response using conventional empirical methods on the model simulated individual-level data, we find much smaller estimates of the Frisch elasticity of labour supply.³

Through a series of experiments aimed at assessing the importance of various modelling assumptions, we conclude that even though the intensive-margin elasticity of aggregate labour supply is quantitatively important, it is almost invariant to changes in the modelling assumptions. The effects of the modelling assumptions on labour supply responses are almost entirely driven by their impact on the extensive margin. When we shut down the non-linear earnings feature in the baseline economy and simulate the employment responses to a small temporary wage change, we find that the fixed costs of work in the baseline economy are sufficiently small that, on their own, they do not affect employment decisions along the extensive margin. However, fixed costs of work matter importantly in the presence of non-linear earnings, accounting for half of the employment response in our baseline economy: when we shut down fixed costs of work in the baseline economy, the extensive margin elasticity to a temporary wage change drops from 1.08 to 0.57. This result underscores that there is a *complementarity* between fixed costs of work and non-linear earnings that *enhances* the aggregate labour supply response to a temporary wage change. We also find that the extensive margin is also crucial for understanding how the labour supply response to a temporary wage change varies across age and education groups. The age profile of the labour supply response in the baseline economy is U shaped. Such a life-cycle pattern is reminiscent of the fact that over the business cycle hours of work and employment fluctuate much more for young and old individuals than for the middle-aged, as discussed in Gomme *et al.* (2005) and Jaimovich and Siu (2009).

Economists assessing the effects of tax policies are interested in estimates of the labour supply response to a permanent (compensated) wage change, as measured by the Hicks elasticity of labour supply. The implicit assumption is that tax revenues are rebated back to the representative consumer so that wage (or tax) changes do not have wealth effects. We find that the Hicks elasticity of aggregate labour supply in the baseline economy is 0.44, which is in line with estimates in the micro-literature.⁴ Our findings suggest that modelling the extensive margin is

2. Keane (2011) surveys 21 of the best known studies that estimate the (Frisch) elasticity of labour supply at the intensive margin, and reports a mean value from the surveyed studies of 0.83. Chetty *et al.* (2011a), in their meta study of 25 papers report a mean value of 0.54 for the (intensive margin) Frisch elasticity of labour supply.

3. The estimates vary substantially depending on the instrument used in the log-hours-log-wage regression (empirical) analysis. Allowing for some measurement error in hours/wages further decrease the estimated elasticities. The fact that the conventional empirical methods lead to estimates that are lower than the elasticity obtained with our ideal macro-experiment is consistent with the views discussed in Keane (2011) and Keane and Rogerson (2012).

4. Keane (2011) surveys 21 studies of the Hicks elasticity of males and reports a mean value of 0.30. The mean value of the Hicks elasticity of aggregate hours across the micro-studies reviewed in Chetty *et al.* (2011a) is 0.59, with

important for an accurate assessment of aggregate labour supply responses to wage (tax) changes and how these responses vary with age. This is because the labour supply responses along the intensive and extensive margins have different signs. While the intensive margin elasticity is 0.55, the extensive margin elasticity is -0.11 . Hence, neglecting the extensive margin response will significantly overstate the aggregate labour supply response to a (compensated) wage (tax) change.⁵ The magnitude of this bias varies across ages, and it is largest late in the life cycle.⁶

Keane (2011) and Keane and Rogerson (2011, 2012) advocate the importance of a research programme that studies how the features studied in our model (*e.g.* life cycle, incomplete markets, preference heterogeneity, non-linear earnings, and the extensive margin) jointly affect labour supply responses. We emphasize that most of the papers in the macro-literature do not model the labour supply decisions of males along the extensive margin (see, for instance, Low, 2005; Domeij and Flodén, 2006; Pijoan-Mas, 2006). It is noteworthy that Keane’s (2011) discussion of the literature on labour supply along the extensive margin refers mostly to females. All of the structural papers he discussed refer to female labour force participation, and they typically use a model period of one year. Perhaps the closest paper to us is French (2005), who incorporates most of the features in our article but models the extensive margin at an annual level because he is mostly interested in studying retirement decisions late in the life cycle.⁷ Finally, our framework abstracts from human capital accumulation, a feature that Imai and Keane (2004) and Keane (2011) have emphasized. Instead, our analysis focuses on different, but complementary, mechanisms.

2. EMPIRICAL FACTS

In this section, we describe the facts on male labour supply over the life cycle, lifetime labour supply, and evidence on non-linear earnings that guide the development of our theory.

2.1. *Labour supply over the life cycle*

We begin our empirical analysis by describing facts on labour supply at the extensive (whether to work or not) and intensive (how much to work) margins. The distinction between the extensive and intensive margins of labour supply depends on the period of time. In our analysis, we will use data from the SIPP, and we will define the time period to be four months (a quadrimester),

substantial variation in the estimated elasticities across studies. Since there are wide confidence intervals associated with each of the point estimates as well as methodological disputes about the validity of some of the studies, Chetty *et al.* (2011a) argue that the estimates should be treated as rough values meant to gauge the order of magnitudes. In this sense, the Hicks elasticity implied by our model appears reasonable.

5. Hours along the intensive margin decrease because of the substitution effect associated with the lower wage rate. On the other hand, the decrease in the (compensated) wage discourages savings over the life cycle, which, in turn, leads to an important increase of the employment rate at old age. This effect accounts for the aggregate negative elasticity of labour supply along the extensive margin.

6. While the Hicks elasticity along the intensive margin is flat over the life cycle, the extensive margin elasticity varies importantly with age. The extensive-margin elasticity for non-college and college individuals aged 55–61 is -0.42 and -0.35 . The corresponding values for the age group 25–34 are -0.07 and -0.05 for non-college and college individuals (*i.e.* about a fifth of those of the oldest age group).

7. Chang and Kim (2006) and Rogerson and Wallenius (2009) are also two well-known papers in the macro-literature that model the extensive margin. The former abstracts from labour supply decisions along the intensive margin, and they do not model preference heterogeneity, life cycle, and tied wage offers (non-linear earnings). Relative to Rogerson and Wallenius (2009), our contribution is to build a theory of aggregation, disciplined with micro-data, with several dimensions of heterogeneity (skills and tastes). Moreover, by modelling non-linear earnings *together* with fixed costs of work, we can focus on how these two features interact in determining labour supply responses along the extensive margin.

implying that variation in labour supply within a quadrimester period is interpreted as changes along the intensive margin while variation in the number of quadrimesters worked would be interpreted as changes in labour supply along the extensive margin.

The SIPP interviews individuals three times a year (once every four months) and provides detailed monthly information on labour market history and income and welfare programme participation. We use this information in order to compute labour market statistics of interest at a high frequency, such as a quadrimester. We use the 1990 SIPP Panel which runs from October 1989 until August 1992. That implies that it spans 8 quadrimesters (32 months).⁸ We restrict the analysis only to men, who are not self-employed, between the ages of 25 and 61. Individuals are classified as non-college if they have either elementary or high school education and as college if they report having completed college education.

Definition of employment and non-employment. In order to classify an individual as either employed or non-employed in a particular quadrimester, we utilize the available information in the SIPP on hours worked at the monthly level.

First, we identify whether an individual in a particular month is: (1) working (*i.e.* reporting positive hours worked), (2) not working (*i.e.* reporting zero hours worked), or (3) a non-respondent, not in the sample. Only months with information about working (*i.e.* either (1) or (2)) can be used in the analysis. Individuals who work as self-employed in a given month fall into category (3) in that particular month. Secondly, if an individual works in one or more months in a given quadrimester, regardless of whether the remaining months fall into categories (2) or (3), then he is classified as employed in that quadrimester. If an individual falls into category (2)—*i.e.* reporting zero hours worked—in each of the four months in a quadrimester, then he is classified as non-employed. All other individuals are dropped from the sample for that quadrimester. Finally, a non-employment spell starts whenever an individual is working in one quadrimester and not in the next one. The non-employment spell ends when the individual records a quadrimester of employment.

In order to discuss life-cycle patterns, below we group individuals in four age groups: 25–34, 35–44, 45–54, and 55–61.

Employment rate and mean hours of work. The quadrimesterly employment rates and mean hours of work by age and education groups are reported in Figures 6 and 9 in the calibration section. The employment patterns are not surprising: college individuals have a higher employment rate, at all ages, than non-college individuals, and there is a pronounced decline in the employment rates late in the life cycle.⁹ Average hours worked are higher for college individuals, and mean hours worked decrease late in the life cycle for both education groups. This is due to both lower employment rates and lower hours per worker, although the extensive margin is quantitatively much more important.

Entry into non-employment spells. Figure 1 reports the fraction of those entering a non-employment spell, by age and education groups. Non-college individuals are much more likely to enter a non-employment spell, and that fraction increases late in the life cycle. Recall that

8. We recomputed all of the empirical facts reported below on the 1993 SIPP data that run from October 1992 until December 1995. The facts are very similar in both SIPP data sets.

9. Note that measures of employment that use information on hours worked in the week preceding the interview will produce slightly lower employment rates than our definition of employment.

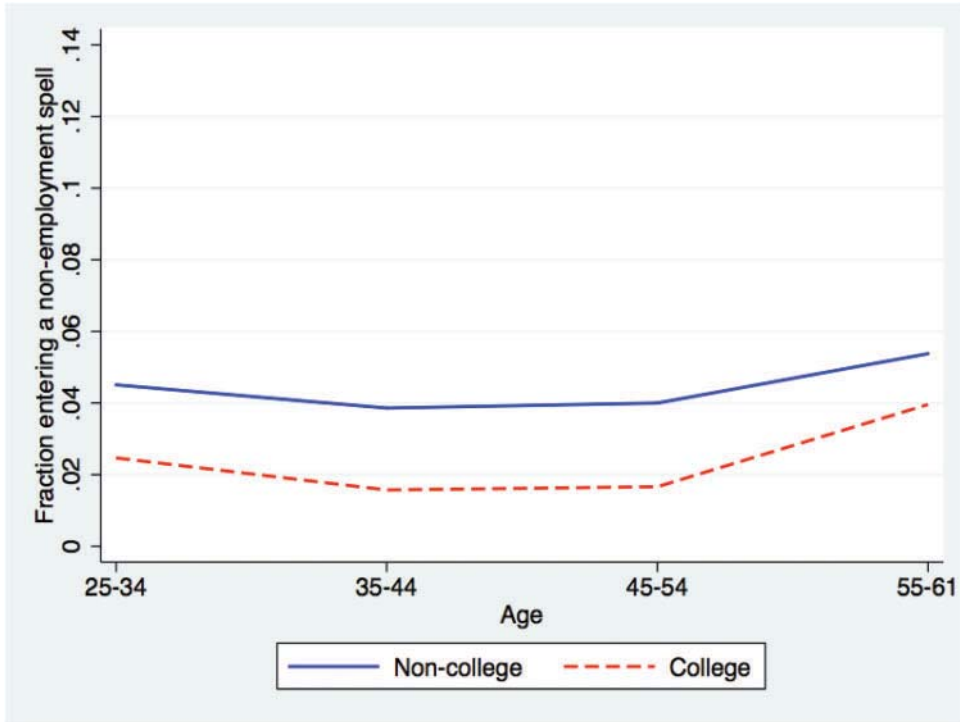


FIGURE 1

Entry rate into non-employment, 1990 SIPP.

Notes: The entry rate into non-employment is measured as the fraction of those employed in quarter t that are non-employed in quarter $t+1$.

individuals are classified as non-employed if they do not work at all throughout the whole four-month period. Therefore, those who use the extensive margin not to work spend a non-trivial amount of time non-employed. Nevertheless, we find very large entry rates into non-employment, at all ages of the life cycle, indicating that individuals consistently use the extensive margin in order to adjust their labour supply: throughout most of the life cycle around 4% (2%) of the employed non-college (college) individuals will find themselves non-employed in the following quarter. These numbers increase to 6% and 4% late in the life cycle for the non-college and the college, respectively.

We use SIPP data on transitions from employment into unemployment in order to differentiate labour market transitions out of employment that might be due to involuntary separations. In order to analyse entry into unemployment, we focus on those individuals who report to have been employed in the previous quarter, but are classified as non-employed in the current quarter according to the criteria specified above. We then use three variables from the 1990 SIPP to identify the subset of individuals from the group of non-employed who are unemployed: (1) number of weeks looking for work, (2) employment status recode for each month, and (3) amount of unemployment compensation and benefits. We consider an individual to be unemployed if he spent any positive time looking for a job or received unemployment benefits and compensation. The resulting quarterly entry rates into unemployment are, as expected, higher for non-college than college individuals. Moreover, for both education groups they are fairly stable over the life cycle: 2.4% for non-college and 1.16% for college. Hence, our

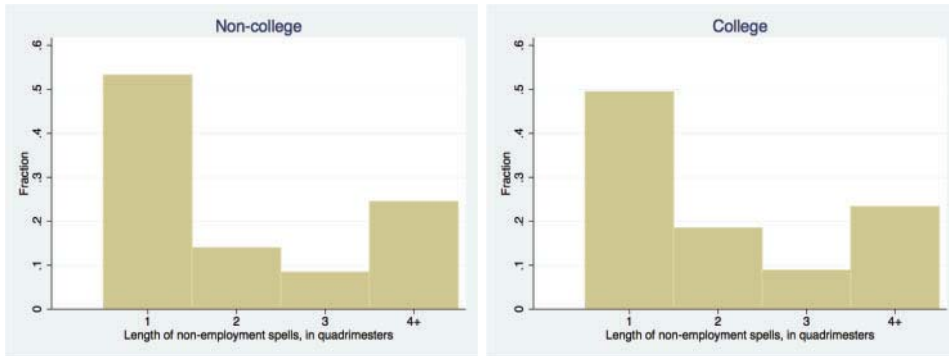


FIGURE 2
Non-employment spells, 1990 SIPP.

findings suggest that about half of the non-employment spells are voluntary. Nonetheless, it is plausible that these estimates underestimate the amount of voluntary separations in the data for two reasons. First, we consider an individual to be non-employed only if the non-employment spell lasts for at least four months, thereby not capturing short non-employment spells. Secondly, our procedure most likely underestimates voluntary separations because an individual who voluntarily separates from a job in a given quadrimester and searches for a job (at some point) during the next quadrimester is counted as unemployed. Hence, if anything, our procedure minimizes the importance of extensive margin responses.

Durations of non-employment spells. Figure 2 displays the histograms of non-employment spells for non-college and college individuals. These histograms provide us with important information regarding the nature of labour supply at the extensive margin. Approximately half of all non-employment spells last for only one quadrimester, and only around 25% of the spells last for more than four quadrimesters (*i.e.* more than 16 months). In general, college individuals tend to spend more periods in non-employment, although as discussed above they are much less likely to enter non-employment spells.

Figure 3 reveals additional insightful information by reporting the histograms of non-employment spells also by age. We observe a pronounced life-cycle pattern, with older individuals spending more periods into non-employment. Nevertheless, even for the old age group of 55–61 year olds, around 50% (60%) of the non-employment spells for non-college (college) individuals last for one or two periods, indicating an active extensive margin of labour supply even at an old age.

2.2. Lifetime labour supply

The dispersion in lifetime labour supply is another useful statistic, which is closely related to the persistence in an individual's labour supply over time. We use the PSID for the period 1968–97 in order to compute all annual statistics. The sample is restricted to men between the ages of 25 and 61. We consider an individual to be high school if he has at most 12 years of education while those with 14 years of education or more are considered to be college graduates. We keep in the sample those born between 1922 and 1962.

Due to the nature of the PSID data set, we do not observe individuals throughout all their life—some of them have already been in the labour market for some time when the survey starts in 1968 while those who enter the labour market in 1968 at the age of 25 are only in their 50s

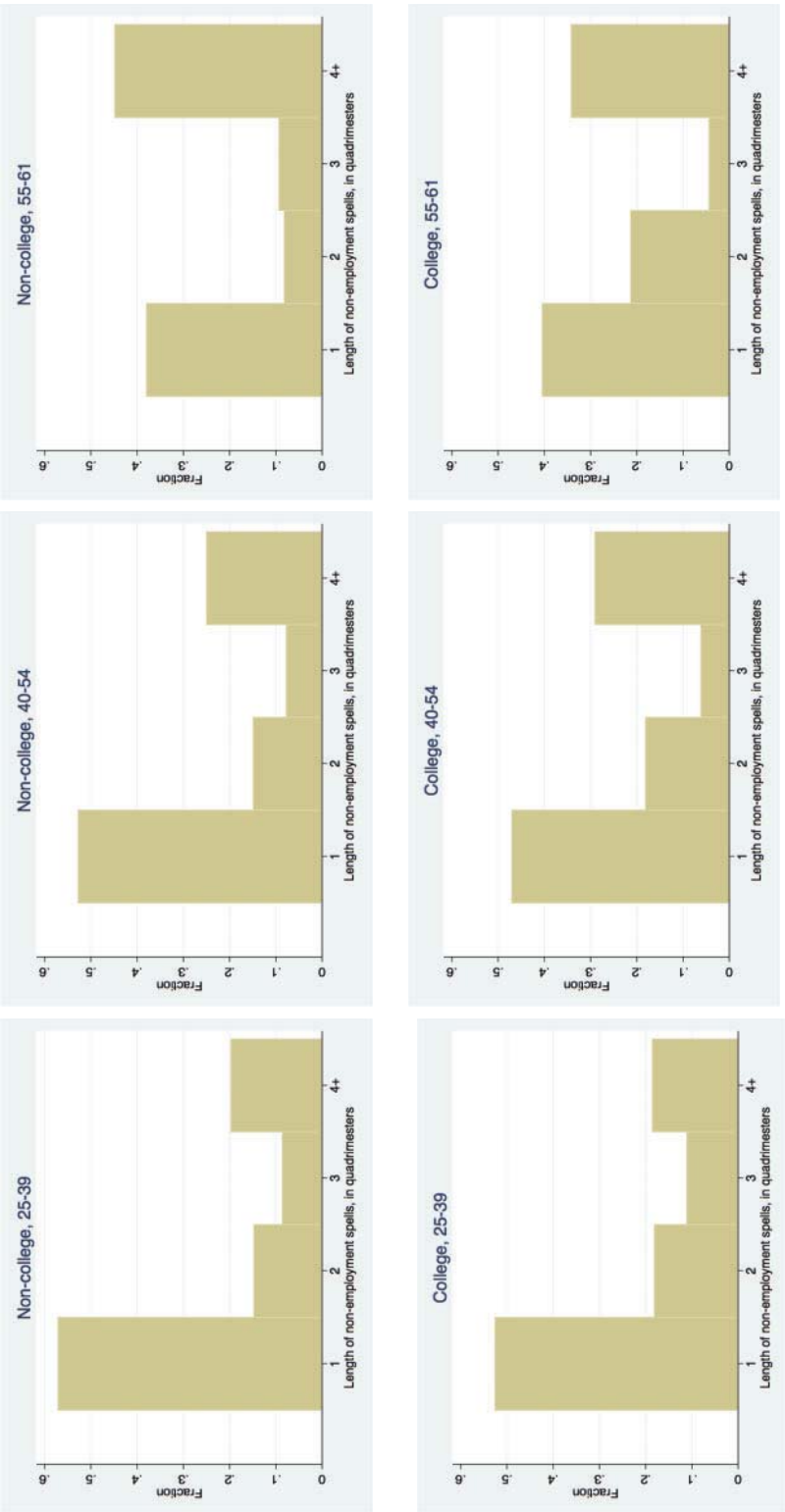


FIGURE 3
Non-employment spells, 1990 SIPP: by age and education.

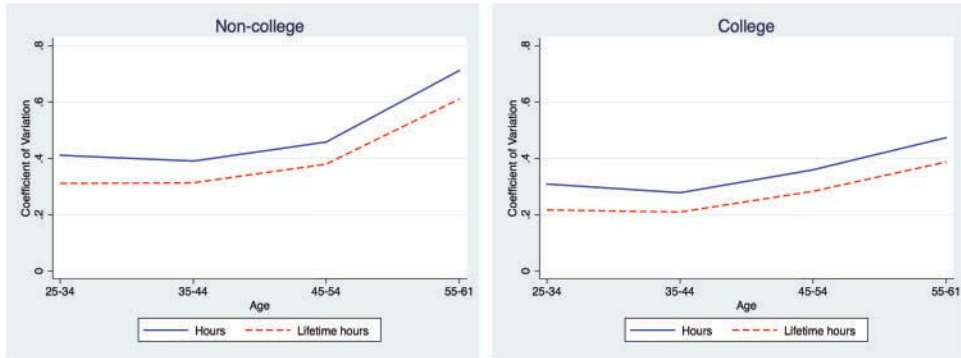


FIGURE 4

Coefficient of variation in hours and lifetime hours, PSID.

Notes: For each 10 year period (e.g. between the ages of 35 and 44), let h_{ij} denote the annual hours worked for individual i who is j years old and let $\tilde{h}_i = (\sum_{j=35}^{44} h_{ij})$ denote the individual i 's lifetime hours during this period. The coefficient of variation in lifetime hours across all i individuals is $CV(\tilde{h}_i)$. The coefficient of variation in hours across all individuals of age j is $CV^j(h_{ij})$. For each age group G , the figure reports the average of $CV^j(h_{ij})$ over all the ages $j \in G$.

in 1997. Nevertheless, we can learn a lot even if we follow individuals for shorter periods. We choose to follow individuals for periods of 10 years at different stages in their life cycle: ages 25–34, 35–44, 45–54, and 55–61. We drop all individuals who have a missing observation during the relevant 10 years and sum the hours worked for each individual during the whole 10 years. Then we compute the coefficient of variation of this 10-year cumulative measure of hours worked and refer to it as a measure of the dispersion in lifetime hours. In addition, we also compute, on exactly the same sample, the coefficient of variation of cross-sectional annual hours and refer to it as a measure of the dispersion in hours.

A comparison of the coefficient of variation of annual hours and lifetime hours reveals information about the persistence of inequality in hours. If the coefficient of variation of lifetime hours is much smaller than that of annual hours, it implies that year-to-year fluctuations in hours, for a given individual, tend to cancel out over time. Hence, when individual hours of work are not persistent across years, inequality in lifetime hours diminishes with the length of the period considered. When inequality in hours across individuals is fixed (or highly persistent), the coefficients of variation in lifetime hours and annual hours are equal (or close to each other).

Figure 4 reports the measured dispersion in cross-sectional and lifetime hours for the four age groups defined above. Non-college individuals display higher dispersions in both cross-sectional and lifetime hours. Further, the analysis provides us with two important findings. First, the dispersion in cumulative hours is quite substantial, indicating that individuals tend to be quite persistent in their labour supply behaviour. Secondly, the dispersion of cumulative hours is smaller than the cross-sectional dispersion at any age in the 25–61 interval. This implies that workers do sometimes change their hours worked.

2.3. Non-linear earnings

In cross-sectional data, there is a positive correlation between hourly wages and hours worked—i.e. those that work more hours also receive higher hourly wages. This result could be due to selection (those who work long hours may be more productive) or could be due to the fact that individuals choose to work less hours during periods of relatively low hourly wages. However, various studies have also documented the existence of an additional mechanism—individuals that

decide (for reasons other than changes in the hourly wage) to work fewer hours do get offered lower hourly wages.¹⁰ For example, Gustman and Steinmeier (1985) document that individuals that enter partial retirement experience—depending on whether they remain in their main job or not—a 15–23% drop in their hourly wages.¹¹ More recently, Aaronson and French (2004) estimate the quantitative importance of this mechanism by exploiting the sharp decline in hours worked for working men at the ages of 62 and 65 in U.S. data. They treat this decline in hours worked for those who work (relative to a quadratic age polynomial in hours worked over the life cycle) as exogenous variation in hours worked, and thus a useful instrument in the empirical analysis, caused by the U.S. social security rules that discourage individuals at those ages from working long hours. They find that men who cut their work-week from 40 to 20 hours experience a 20–25% decline in their hourly wage. In order to control for the fact that workers that reduce their working hours might be also changing the type of work they do on the job, either by switching employers or by switching their occupation within the same employer, they further restrict the analysis to workers that remain in the same job and/or same occupation. Finally, since the various data sets used in the analysis are all longitudinal and follow individuals over time, the estimation controls for individual fixed effects and thus for unobserved quality and productivity differences between those that work long and short number of hours.¹²

3. MODEL

We develop a life-cycle theory of labour supply of individuals along the intensive and extensive margins. The model abstracts from the labour supply decisions of women and analyses only males. We consider a small open economy facing a fixed interest rate.

3.1. *Population, preferences, and endowments*

The economy is populated by overlapping generations of individuals who start their lives at age 25, face uncertain lifetimes, and live, at most, J periods. They differ in terms of their education (college versus non-college), labour productivity, and taste for leisure. The college decision is exogenous, and the education type of an individual determines the stochastic processes driving the mortality, taste, and labour productivity shocks.

Preferences. The date- t utility function takes the form

$$u_t = u(c_t, l_t) = (1 - \varphi) \ln c_t + \varphi \frac{l_t^{1-\sigma}}{1-\sigma} - I_{(l_t < 1)} F, \quad (1)$$

where c_t is consumption and l_t denotes leisure. Individuals are heterogeneous in their taste for leisure $\varphi \in (0, 1)$, which evolves stochastically over time. Individuals who work ($l_t < 1$) face a fixed disutility cost of work (F). The utility function is consistent with balanced growth—this assumption allows the theory to be consistent with the fact that there are large permanent

10. See Gustman and Steinmeier (1985) and Aaronson and French (2004) for references to the literature that has studied this issue.

11. Gustman and Steinmeier (1986) argue that a structural model of retirement with such a feature matches very well the labour supply patterns after the age of 60.

12. Moffitt (1984), using taste shifters, such as children, as instruments in a wage on hours regression, and Keane and Wolpin (2001), using a structural model, also find evidence of tied wage-hours offers.

differences in labour productivities across individuals (heterogeneity in fixed effects) but not in their lifetime labour supply.¹³

Individuals maximize lifetime expected utility

$$E \sum_{t=1}^J \beta^t \pi_t u(c_t, l_t), \quad (2)$$

where π_t denotes the probability that the individual survives to age t and E denotes the expectation operator. Each period, as described below, individuals face mortality shocks, labour market risk (job separation and job finding risk), shocks to the taste for leisure (φ), and labour productivity risk (z).

Heterogeneity in preferences. For each individual, it is assumed that taste shocks follow the stochastic process

$$\varphi_{it} = \varphi_i \times \varphi_t,$$

where $\varphi_i \sim N(\mu_\varphi, \sigma_\varphi^2)$ is an individual fixed effect determined at birth. The term φ_t is an individual-specific shock that represents a stochastic deviation from the mean value φ_i , which follows a first-order autoregressive process.

$$\begin{aligned} \varphi_t &= \rho_\varphi \varphi_{t-1} + \eta_{\varphi_t}, \quad \eta_{\varphi_t} \sim N\left(1 - \rho_\varphi, \sigma_{\eta_\varphi}^2\right), \\ \varphi_0 &\sim N\left(1, \frac{\sigma_{\eta_\varphi}^2}{1 - \rho_\varphi^2}\right), \end{aligned} \quad (3)$$

where ρ_φ denotes the persistence of the shock on preferences for leisure and η_{φ_t} is the innovation at age t . When individuals enter the model economy, the initial seed for the autoregressive process (φ_0) is drawn from the invariant distribution of preference shocks. This assumption implies that the distribution of shocks does not vary over the life cycle, so that the life-cycle patterns of labour supply implied by the theory are not due to variation in preferences over the life cycle. While the parameters ($\mu_\varphi, \sigma_\varphi^2, \rho_\varphi, \sigma_{\eta_\varphi}^2$) vary across education types, this is omitted to simplify the notation.

Labour services. An individual's time endowment in each period is one. The amount of time that can be allocated to work is $h_j = 1 - l_j$. It is assumed that the mapping from hours of work to labour services is non-linear, as in Hornstein and Prescott (1993), French (2005), and Rogerson and Wallenius (2009). The idea is that h hours of work map into h^θ units of labour services, with $\theta \geq 1$. The case $\theta = 1$ corresponds to the standard model that assumes a linear mapping from hours to labour services. The case $\theta > 1$ gives rise to a model economy in which earnings are a non-linear function of hours of work. Earnings of an individual with labour productivity z and working hours h are

$$w(z, h) = zh^\theta. \quad (4)$$

13. To illustrate this point, for each cohort and education group we divide individuals into high and low productivity types. We compute each individual's mean wage over the age of 30–45 and classify them into high and low types depending on whether their mean wages are above or below the median wage in their cohort-education category. We then compute mean hours worked for high and low types and find that there are virtually no differences in labour supply. Focusing on the age group 30–45 and non-college individuals, the average hours worked across all cohorts is 2143 for type 1 individuals and 2166 for the type 2. For individuals with college education, average hours are 2269 and 2271 for type 1 and type 2, respectively.

The calibration of the model will rely on estimates by Aaronson and French (2004) to pin down the value of θ . In the Online Appendix A, we use the production technology proposed by Hornstein and Prescott (1993) to show that it gives rise to a competitive equilibrium in which earnings follow the functional form specified in (4).

Heterogeneity in labour productivity. For each education group, labour productivity z_t is assumed to change stochastically over the life cycle according to:

$$\ln(z_{it}) = \mathbf{x}_t \kappa + \alpha_i + u_t, \quad (5)$$

where z_{it} denotes labour productivity of individual i at age t , \mathbf{x}_t is a quartic polynomial in age, κ is a vector of coefficients, $\alpha_i \sim N\left(-\frac{\sigma_\alpha^2}{2}, \sigma_\alpha^2\right)$ is a fixed effect determined at birth, and u_t is a persistent productivity shock. Each period, it is assumed that with probability two-thirds the productivity shock remains constant $u_t = u_{t-1}$ and with probability one-third it follows a first-order autoregression:

$$u_t = \rho_u u_{t-1} + \eta_{u_t}, \quad \eta_{u_t} \sim N\left(-\frac{\sigma_{\eta_u}^2}{2}, \sigma_{\eta_u}^2\right), \quad u_0 = 0. \quad (6)$$

While the parameters $\Omega = (\kappa, \sigma_\alpha^2, \rho_u, \sigma_{\eta_u}^2)$ vary across education types, this is omitted to simplify the notation. At age 62, labour productivity becomes zero so that all individuals are retired by this age.

Labour market frictions. Individuals start each period in three possible labour market states: employed (e), non-employed (n), and unemployed (u). Employed individuals (e) at the beginning of the period are those that have a job offer at that point, and face the decision of whether to work or not. The main distinction between the last two states (n and u) is that unemployment insurance benefits b_u are assumed to be paid only to the unemployed. The transitions in labour market status across periods depend on the labour supply decisions of individuals (whether to work or not) and on two labour market shocks, an exogenous job separation shock (δ) and a job finding rate (p), as described in Table 1. Individuals that start the period as employed and choose to work ($h > 0$) face at the end of the period a job separation shock: with probability δ they are exogenously separated and start next period in the unemployment state, or with probability $(1 - \delta)$ they do not suffer a job separation and start next period as employed. Individuals that start the period as employed and choose not to work ($h = 0$) find a job at the end of the period with probability p so that next period they start as either employed (with probability p) or non-employed (with probability $1 - p$). Since unemployed individuals find a job with probability p , they transit to the employment state with such probability. Unemployed individuals that do not find a job (an event with probability $1 - p$) transit next period to the non-employment state. Since unemployed individuals transit into either employment (with probability p) or non-employment (with probability $1 - p$), individuals can only collect unemployment insurance during the period they were exogenously separated from their job. The parameters δ and p do not vary with age over the life cycle, but they do vary across the two education groups. We also assumed that both unemployed and non-employed individuals face the same job finding rate p .

3.2. Government

The government taxes consumption, capital income, and labour income. It is assumed that consumption and capital income are taxed at flat rates (τ_c, τ_k). Earnings y are taxed according to a

TABLE 1
Labour market transitions

<i>s</i> and hours	<i>s'</i>		
	<i>e</i>	<i>u</i>	<i>n</i>
<i>e</i> , <i>h</i> > 0	1 − δ	δ	0
<i>e</i> , <i>h</i> = 0	<i>p</i>	0	1 − <i>p</i>
<i>u</i>	<i>p</i>	0	1 − <i>p</i>
<i>n</i>	<i>p</i>	0	1 − <i>p</i>

progressive tax schedule $T(y)$. The tax revenue is used to finance government expenditures, which are assumed not to provide utility to individuals or to enter in an additively separable fashion in the utility function. The government administers an unemployment insurance scheme that provides unemployment benefits for one period to those individuals who were laid off in the current period. Unemployment benefits are set to a fixed proportion of individual's potential earnings, computed as the earnings the individual would have made had he worked full time—40% of available time—during the quadrimester.

The government also administers a pay-as-you-go social security system. We model a stylized representation of the U.S. social security system. Pension benefits in the U.S. are a function of the Average Indexed Monthly Earnings (AIME) over the 35 highest earnings years. Modelling in detail how pension benefits depend on the history of earnings of individuals requires the modelling of a state variable that summarizes average past earnings—an approach which substantially complicates the computational task in a model which is rich in many other dimensions. To simplify the computational procedure, we follow a large literature (see, for instance, (Low *et al.*, 2010) and (Kaplan, 2012a)) in assuming that pension benefits are a function of individual's fixed characteristics: education and the fixed effect in labour market productivity. We assume that the government collects a payroll tax (τ_{ss}) to finance social security outlays.

3.3. Credit markets

Individuals can insure mortality risk in fair annuity markets.¹⁴ Denoting by R the gross interest rate net of capital income taxes τ_k , the gross interest rate faced by an individual j years old with education e is given by

$$R_j^e = 1 + \left(\frac{1+r}{\pi_j^e} - 1 \right) (1 - \tau_k), \quad (7)$$

where π_j^e is the conditional probability that an age $j - 1$ individual with education e survives to age j . We assume that individuals can borrow up to an exogenous fixed limit ($a' \geq \underline{a}$) except for the last period of life.¹⁵ The fact that the interest rate is adjusted by the individual's mortality rate is derived from a zero-profit condition in the problem of financial intermediaries competing for deposits ((Erosa *et al.*, 2012)). At age T we impose a non-negative constraint on savings ($a_{T+1} \geq 0$) capturing the fact that financial intermediaries are not willing to make loans to individuals that will be dead next period with probability 1. Note that there are no accidental bequests due to

14. This is done to avoid taking a stand on how accidental bequests are distributed across individuals. Note that the taxation of capital income makes the after tax return on annuities actuarially unfair.

15. Keane and Wolpin (2001), in a structural model of educational attainment, find that borrowing constraints for youth are fairly tight. Keane and Wolpin (2001) also provide an overview of the literature on the importance of borrowing constraints on educational attainment and life-cycle consumption.

the presence of annuity markets. Nonetheless, from the perspective of households annuities are unfair because of the taxation of capital income.

3.4. *The individual's problem*

We use recursive language to describe the problem of an individual. To simplify the notation, we abstract from the fact that the education type of an individual determines his labour productivity process, mortality risk, and labour market shocks. Then, the state of an individual is given by his age j , assets a , and earnings shock z , taste shock φ , and labour market status s . At the beginning of the period, prior to the labour supply decision, individuals can be in one of three labour market states s : employed (e), non-employed (n), and unemployed (u). As discussed above, the only difference between being in the labour market states (n) and (u) is that in the latter individuals can collect unemployment insurance benefits.

Earnings can take one of several forms: labour earnings, unemployment insurance b_u , and pension benefits paid by the social security system b_s . Earnings are assumed to be taxed progressively according to a progressive tax schedule $T(y)$. The budget constraint is then given by

$$V_j(a, z, \varphi, s) = \max_{\{c, h, a'\}} \{u(c, l, \varphi) + \beta \pi_{j+1} E[V_{j+1}(a', z', \varphi', s'(h, s))]\} \quad (8)$$

s. t.

$$a' = y(z, h, s) - T(y(z, h, s)) + R_j a - c(1 + \tau^c),$$

$$a' \geq \underline{a},$$

where the expectation operator is taken over productivity, taste, and labour market shocks. The labour market state next period, s' , is a stochastic function of the current labour market state, s , and current working hours h , as described in Table 1. We impose mandatory retirement at the age of 62 at which point individuals start collecting pension benefits. The only source of uncertainty faced by retired individuals is in their preference shock (φ) and mortality shocks.

3.5. *Discussion on non-linear earnings and labour supply decisions*

In the Online Appendix B, we show the following properties of labour supply decisions in a simplified version of our economy with non-linear earnings:

- Intensive margin: Hours of work along the intensive margin are bounded away from zero. They are decreasing in non-labour income (x) and taste for leisure (φ), and they are increasing in labour productivity (z). The Frisch elasticity of leisure is determined by the curvature of the utility function on leisure (σ), and it is not affected by the assumption of non-linear earnings. The Frisch elasticity of labour supply along the intensive margin decreases with hours of work and is given by $\eta_h = \frac{1}{\sigma}(1 - h)/h$.
- Extensive margin: Non-linear earnings penalize individuals working low hours because marginal earnings are close to zero when hours are low. Non-linear earnings encourages individuals to work either long hours or not at all. This trade-off makes the extensive margin of labour supply decisions more prominent, even in the absence of fixed costs of work $F=0$. Ceteris paribus, individuals that choose not to work are characterized by either a high value of non-labour income (x), taste for leisure (φ), or low productivity (z). The aggregate Frisch elasticity of labour supply along the extensive margin is determined

by preference parameters (φ, σ) , technology parameters (θ) , and the joint distribution of individuals characteristics (x, z, φ) in the population.

- Homotheticity: A doubling of labour earnings (z) and non-labour income (x) has no consequences for labour supply decisions along the extensive and intensive margins.

4. CALIBRATION

We divide the parameters of the model economy in two groups. The first group includes parameters that are pinned down without simulating the model economy (such as, mortality rates and tax rates). The second group is composed of parameters that are calibrated by simulating the model economy.

4.1. *Parameters calibrated without simulating the model economy*

The model period is set to one quadrimester (four months). The model economy is solved in partial equilibrium for a fixed interest rate. The quadrimesterly interest rate is chosen so that the implied annual rate of return on capital (net of depreciation) is 4%.

The intertemporal elasticity of substitution of leisure. In the calibrated baseline economy, we choose $\sigma = 2.0$ that implies an intertemporal elasticity of leisure of 0.5. We also report the results for an alternatively calibrated economy with $\sigma = 3$ (an intertemporal elasticity of leisure of one-third).

Non-linear earnings. The hourly wage in our theory satisfies

$$w_h(z, h) = \frac{zh^\theta}{h} = zh^{\theta-1}. \quad (9)$$

Note that the elasticity of the wage rate to a change in hours of work is given by $(\theta - 1)$. As we discussed in Section 2.3, in an empirical study Aaronson and French (2004) estimate this elasticity to be around 0.40. This estimate implies that a full time (40 hours a week) worker earns an hourly wage 20% higher than a part time (20 hours) worker. We thus set $\theta = 1.4$.

Tax rates, social security, and unemployment insurance benefits. The tax rate on consumption τ_c is set at 0.075 as in McDaniel (2007). Following Domeij and Heathcote (2004), taxes on capital income are set to $\tau_k = 0.40$. The social security tax rate is set to $\tau_{ss} = 0.153$, and the cap \bar{y} on social security taxation is fixed at 2.47 of average earnings in the economy (\bar{y}). Taxes on earnings (non-capital income) are set according to a progressive tax schedule with 5 tax brackets. Following French (2005), we parameterize the tax schedule using data on U.S. federal and state income taxes for household heads (with the standard deduction) in 1990 in the state of Rhode Island, which is a fairly representative state in terms of its income tax system. The tax brackets are defined by the following thresholds expressed as multiples of the average earnings (\bar{y}) in the economy: 0.10, 0.16, 0.63, and 1.93. The corresponding tax rates for the five tax brackets are 0%, 13.2%, 17.9%, 32.9%, and 36.9%.

As discussed before, for computational simplicity we follow a vast literature in modelling a stylized representation of the U.S. social security system. For each education level, we assume two different values for average lifetime earnings (one value for each of the two possible fixed effects on labour productivity). Then the average earnings for each of the four productivity types are computed according to the benefit formula of the U.S. social security system (the U.S. social

security benefit formula is a function of the Primary Insurance Amount and has two bend points at 0.2 and 1.24 of the average earnings in the economy).

The unemployment insurance benefit is set to 40% of the potential earnings of the unemployed worker, where potential earnings are computed assuming a work week of 40% of available time.

Mortality rates. The mortality risk for college and non-college individuals is taken from Bhattacharya and Lakdawalla (2006).

Exogenous job separation rates into unemployment. Based on the empirical analysis in Section 2.1, we set the exogenous separation rates into unemployment (δ) in a quadrimester to 2.4% for non-college and 1.16% for college.

4.2. *Parameters calibrated by simulating the model economy*

Rather than simulating the model to calibrate all the parameters at once, we find it convenient to partition the parameters in two subgroups and to follow an iterative procedure with two nested loops. The parameters in each of the two subgroups are calibrated in two separate loops in order to diminish the dimensionality of the calibration. In the inner loop, for fixed values of the parameters calibrated in the outer loop, we calibrate the parameters determining the stochastic process on taste for leisure, the job finding rate, and the fixed utility cost of work. Given the parameters obtained in the inner loop, the outer loop calibrates the parameters determining the labour productivity process, discount rate, average earnings in the economy (used to define tax schedules), average lifetime earnings for the four fixed productivity types (used to compute pension benefits for the two fixed productivity types in each of the two education groups), and the exogenous borrowing limit.

4.2.1. Inner loop. The inner loop pins down the parameters determining the heterogeneity in the taste for leisure (φ), the fixed utility cost of work (F), and the job finding rate (p). Recall that taste for leisure (φ) depends on an individual fixed effect (φ_i) and a stochastic shock (φ_t). The fixed effect is drawn from a normal distribution, $\varphi_i \sim N(\mu_\varphi, \sigma_\varphi^2)$. The stochastic shock (φ_t) represents a deviation from the mean value and follows a first-order autoregressive process with persistence ρ_φ and innovation η_φ (see equation (3)). The innovation is drawn from a normal distribution with variance $\sigma_{\eta_\varphi}^2$. The initial value of φ_t (at age 25) is drawn from the invariant distribution so that the cross-sectional distribution of taste shocks does not vary over the life cycle.¹⁶ Hence, we need to pin down, for each education group, six parameters: four values determining the heterogeneity in taste for leisure ($\mu_\varphi, \sigma_\varphi^2, \rho_\varphi, \sigma_{\eta_\varphi}^2$) plus the fixed utility cost of leisure F and the job finding rate p . The values of these parameters are obtained by minimizing the distance (square of the sum of deviations) between the model and the SIPP data on the following statistics:

1. Average employment rates in a quadrimester for four age groups (25–34, 35–44, 45–54, and 55–61), as computed in Section 2.1 from the SIPP.
2. The probability of entering a non-employment spell in a quadrimester for four age groups (25–34, 35–44, 45–54, and 55–61), as computed in Section 2.1 from the SIPP.

¹⁶ The continuous stochastic process in the model for the taste for leisure is approximated with a Tauchen procedure, with nine possible values (three for the fixed effects and three for the persistent taste shock).

3. The fraction of all non-employment spells lasting one, two, three, or more than three quadrimesters, as computed in Section 2.1 from the SIPP.
4. The coefficient of variation of lifetime hours for workers between the ages of 35 and 44, as computed in Section 2.2 from the PSID: 0.35 for non-college and 0.23 for college individuals, respectively.
5. Average mean hours of work in a quadrimester among prime-age males. Following Osuna and Ríos-Rull (2003) and Prescott (2004), the time endowment is set at 5200 hours a year (100 hours per week). Using data from the 1990 SIPP on prime-age males with positive hours of work, 35–50 years old, we obtain that non-college and college individuals work, on average, about 41.2% and 43.5% of their time endowments in a quadrimester.

Therefore, we use 14 moments from the data in order to parameterize the six parameters in this stage of the calibration process. Even though there is no one-to-one mapping between the parameters and the moments in the data, it is possible to provide intuition as to which moments help identify the model parameters. The job finding rate (p) is identified from the fraction of non-employment spells lasting one, two, three, or more quadrimesters. However, most powerful among these is the fraction of non-employment spells lasting one quadrimester—individuals that experience an exogenous job separation would usually choose to immediately go back to work after one quadrimester, and a job finding rate of $p < 1$ is the only friction preventing that. The fixed utility cost of work (F) is identified from employment rates over the life cycle—higher fixed utility costs of work lead to lower employment rates. The mean of the fixed effect in the taste for leisure (μ_φ) is identified from the mean hours of work in a quadrimester while the variance in the fixed effect in the taste for leisure (σ_φ^2) is identified from the coefficient of variation in the lifetime hours worked—a higher dispersion in the lifetime hours implies a higher variance in the taste for leisure fixed effect. Finally, the persistence and the innovation to the taste for leisure stochastic shock ($\rho_\varphi, \sigma_{\eta_\varphi}^2$) are identified from the entry rates into non-employment (especially those that are not due to exogenous job separations) and the fraction of non-employment spells lasting two, three, or more quadrimesters.

4.2.2. Outer loop. The outer loop pins down the parameters governing the labour productivity process, discount rate, average earnings in the economy (used to define tax schedules), average lifetime earnings for four productivity types (in order to compute social security benefits), and the exogenous borrowing limit. The labour productivity process is calibrated, through an iterative procedure that ensures that the model is consistent with data moments on hourly wages over the life cycle. Since the SIPP is longitudinally fairly short to allow us to estimate the stochastic process for wages, we use the PSID for this purpose. In calibrating a *quadrimesterly* stochastic process on labour productivity, one difficulty arises from the fact that the PSID only reports earnings and hours of work at an annual frequency. Moreover, we need to consider that the data only report wages for individuals that work. To deal with these problems, we follow an iterative procedure:

1. Estimate a wage profile and wage process for college and non-college workers on annual data from the PSID data. In particular, we specify the following annual process for log hourly wages:

$$\ln(\tilde{w}_h)_{ij} = \tilde{\mathbf{x}}_j \tilde{\kappa} + \tilde{\alpha}_i + \tilde{u}_j + \tilde{\lambda}_j, \quad (10)$$

where $\ln(\tilde{w}_h)_{ij}$ represents the observed annual log hourly wage of individual i at age j in the PSID data, $\tilde{\mathbf{x}}_j$ is a quartic polynomial in age, $\tilde{\kappa}$ is a vector of coefficients, $\tilde{\alpha}_i \sim N\left(-\frac{\sigma_\alpha^2}{2}, \sigma_\alpha^2\right)$

is a fixed effect determined at birth, $\tilde{\lambda}_j \sim N\left(-\frac{\sigma_z^2}{2}, \sigma_{\lambda_j}^2\right)$ is an idiosyncratic transitory shock which is interpreted as measurement error, and \tilde{u}_j follows a first-order autoregressive process:¹⁷

$$\tilde{u}_j = \rho_{\tilde{u}} \tilde{u}_{j-1} + \tilde{\eta}_{u_j}, \quad \tilde{\eta}_{u_j} \sim N\left(-\frac{\sigma_{\tilde{\eta}_u}^2}{2}, \sigma_{\tilde{\eta}_u}^2\right), \quad \tilde{u}_0 = 0. \quad (11)$$

2. Feed a *quadrimesterly* labour productivity process $\Omega = (\kappa, \rho_u, \sigma_\alpha^2, \sigma_{\eta_u}^2)$ into the model economy as specified in Section 3.1.¹⁸
3. Simulate the model economy to obtain *quadrimesterly* data on employment, hours of work, and earnings.
4. Aggregate the *quadrimesterly* data to an annual period.
5. Estimate an annual hourly wage profile and hourly wage stochastic process for college and non-college workers in the model generated data.
6. Feed a new *quadrimesterly* labour productivity process (go back to step 2), until the “same” annual wage profile and stochastic wage process is obtained in the model and in the data.

Following Kaplan and Violante (2010), the discount factor β is chosen to match an asset to income ratio of 2.5. This is the wealth to income ratio when the top 5% of households in the wealth distribution are excluded from the Survey of Consumer Finances (SCF). The reason for excluding the richest households in computing an aggregate wealth to income ratio is that the PSID under-samples the top of the wealth distribution. The borrowing limit (\underline{a}) is pinned down so that the model is consistent with the consumption growth rate between ages 25 and 55 of about 25% (Fernández-Villaverde and Krueger, 2007). We assume a common value of β and \underline{a} for the two education groups since we target aggregate data, across the two education groups, on the asset to income ratio and consumption growth over the life cycle.

4.3. Calibration results

4.3.1. Outer loop. Our calibration procedure implies a value for the discount factor of $\beta = 0.983$ and a borrowing constraint equivalent to 40% of average annual earnings in the economy. The latter estimate is similar to the one obtained by Kaplan (2012a). As argued in Kaplan (2012a), if borrowing were not allowed consumption growth predicted by the theory would be counterfactually high. The calibration implies a capital to output ratio of 2.6 and a mean consumption growth over the life cycle of 24%, which are close to the targets of 2.5 and 25% respectively. Table 2 reports the values of the parameters driving the stochastic process on labour productivity in the baseline economy. These parameters were determined using the iterative procedure described above (see Section 4.2.2). The stochastic process on wages estimated on the model data and on the PSID data deliver similar estimates. Moreover, Figure 5 shows that the model mimics well the life-cycle profile of wages.

17. We follow Meghir and Pistaferri (2004) and Low *et al.* (2010) in interpreting the purely iid component in residual log wages (or annual earnings) as measurement error. Meghir and Pistaferri (2011) provide an overview of the literature and the available specifications of the stochastic processes of wages (and earnings).

18. A Tauchen procedure is used to approximate the stochastic process with a finite number of realizations of the shock. We use two values for the fixed effect and 20 values for the persistent shock.

TABLE 2
Calibration of the stochastic process for wages

Non-College				
Parameter	Value	Moment	Data	Model
σ_{α}^2	0.094	$\sigma_{\tilde{\alpha}}^2$	0.094	0.097
ρ_u	0.949	$\rho_{\tilde{u}}$	0.941	0.940
$\sigma_{\eta u}^2$	0.014	$\sigma_{\tilde{\eta u}}^2$	0.019	0.019

College				
Parameter	Value	Moment	Data	Model
σ_{α}^2	0.065	$\sigma_{\tilde{\alpha}}^2$	0.075	0.074
ρ_u	0.982	$\rho_{\tilde{u}}$	0.968	0.977
$\sigma_{\eta u}^2$	0.016	$\sigma_{\tilde{\eta u}}^2$	0.019	0.020

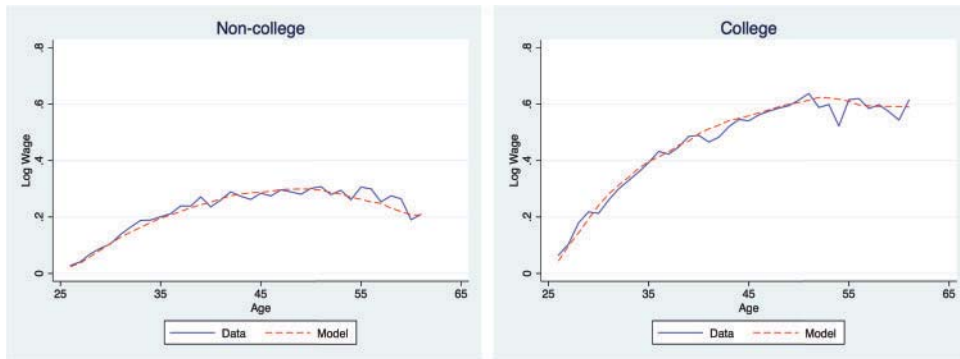


FIGURE 5
Deterministic life-cycle wage profile, PSID and model.

TABLE 3
Calibration of parameters in the inner loop

Variable	Parameter	Non-college	College
Prob. of finding a job	p	0.69	0.59
Mean of fixed effect of taste for leisure	μ_{φ_i}	0.60	0.55
Std. dev. of fixed effect of taste for leisure	σ_{φ_i}	0.12	0.11
Pers. of transitory shock of taste for leisure	ρ_{φ_t}	0.73	0.21
Std. dev. of transitory shock of taste for leisure	$\sigma_{\eta_{\varphi}}$	0.006	0.008
Fixed cost of work (% consumption loss)	\bar{F}	0.085	0.105

4.3.2. Inner loop. Table 3 reports the parameters obtained in the inner loop. The calibration implies that individuals from both education groups face a job finding rate below one (0.69 and 0.59 for non-college and college) and substantial fixed (utility) costs of work. To make sense of the importance of the fixed costs of work, it is convenient to rewrite the utility

function as follows:

$$\begin{aligned}
u(c_t, l_t) &= (1 - \varphi) \ln c_t + \varphi \frac{l_t^{1-\sigma}}{1-\sigma} - I_{(l_t < 1)} F, \\
&= (1 - \varphi) \ln \left(\frac{c_t}{1 + \hat{F}} \right) + \varphi \frac{l_t^{1-\sigma}}{1-\sigma} \\
\text{where } \hat{F} &\equiv \exp \left(\frac{F}{1-\varphi} \right) - 1 \text{ if } l_t < 1, \text{ and equal to 0 otherwise.}
\end{aligned}$$

The calibration implies the utility cost of work is equivalent to losses in consumption of 8.5% and 10.5% for non-college and college individuals, respectively. In Section 5.1.2, we argue that these numbers are consistent with the evidence from Aguiar and Hurst (2013) on work-related expenditures. The calibration implies substantial heterogeneity in the preference shocks for the non-college and college individuals, with a coefficient of variation of around 0.20 for both education groups.¹⁹

4.3.3. Other moments targeted in the calibration. We now show how the baseline model economy matches the moments targeted in the calibration. We emphasize that, except for the age-profile of labour productivity, all the parameters in the model do not vary with age. In particular, our calibration assumes that the job separation and job finding rates, the fixed cost of work, and the distribution of preference shocks are constant over the life cycle. Hence, the life-cycle patterns shown below are the result of the model features driving life-cycle behaviour: a finite time horizon, asset accumulation for retirement, precautionary savings, and the age profile of wages.

Employment rates. The baseline economy is consistent with the fact that the age profile of employment is roughly flat early in the life cycle for both education groups (Figure 6). The model tends to over-predict employment rates, though not by much. The baseline economy captures the fact that for non-college individuals employment starts declining at age 45 and that this decline accelerates substantially after age 55. Moreover, consistent with the data, the decline in the employment rate for college individuals starts at an older age than for non-college. Overall, the model captures quite well, qualitatively and quantitatively, the decline in employment rates late in the life cycle.

Hours worked along the intensive margin and lifetime labour supply. The model economy matches the targets of 0.412 and 0.435 for the average hours of work for prime-aged college and non-college individuals (aged 35–50), respectively. Matching these targets requires that the weight on leisure in the utility function be higher for non-college than for college individuals (0.60 versus 0.55).

The model is also consistent with the heterogeneity in lifetime labour supply in the PSID data. The calibration targeted the coefficient of variation of lifetime hours of work for the 35–44 age group, which takes the value of 0.35 for non-college and 0.23 for college individuals, respectively (see Section 2.2). To match this fact, the calibration requires substantial heterogeneity in the

19. The patterns in the preference heterogeneity are consistent with the analysis in Heathcote *et al.* (2014) who, using different identifying restrictions, reach similar conclusions.

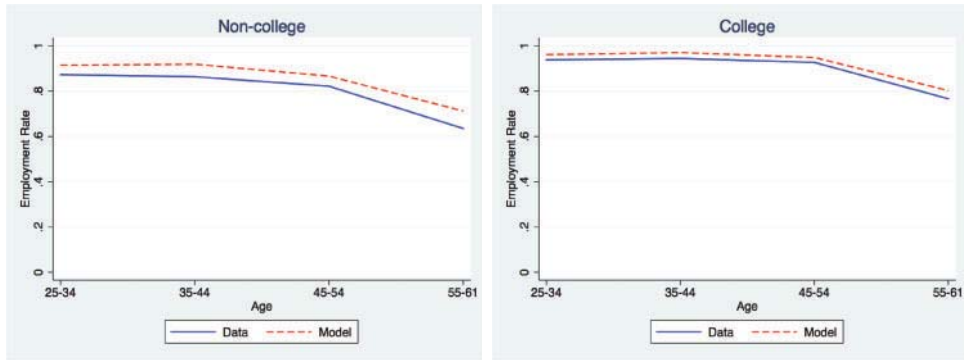


FIGURE 6
Employment rate, 1990 SIPP and model.

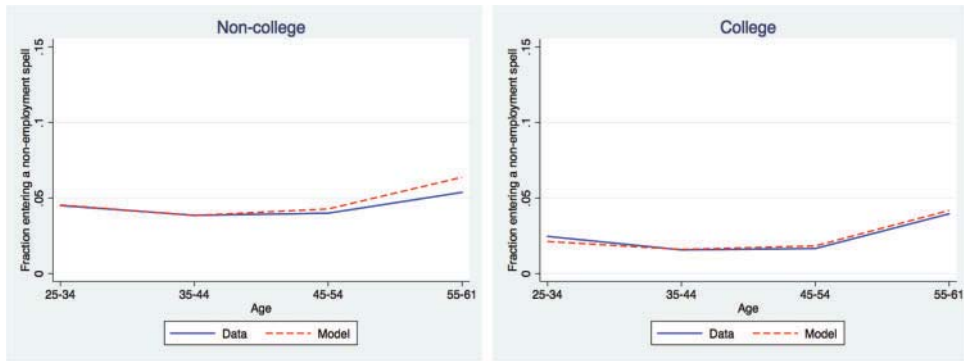


FIGURE 7
Entry rate into non-employment, 1990 SIPP and model.

Notes: The entry rate into non-employment is measured as the fraction of those employed in quadrimester t that are non-employed in quadrimester $t + 1$.

taste for leisure for both education groups, with a bigger variance for non-college than college individuals.

Entry into non-employment spells. The calibration targets the quadrimesterly hazard rates into non-employment over the life cycle. Note that the transition rates from employment into non-employment provide information on the variation of labour supply decisions along the extensive margin. Hence, this target is important for disciplining the predictions of the theory for how the aggregate employment rate responds to temporary wage shocks. Figure 7 shows that the model matches quite well the hazard rates into non-employment. The baseline economy matches well the fact that the hazard rates have a U shape over the life cycle and are higher for non-college than college individuals. We emphasize that the life-cycle profile for hazard rates are not due to variation in preference shocks over the life cycle. As we shall discuss later, allowing for (positive) borrowing and modelling preference heterogeneity is important for the decline in hazard rates early in the life cycle. The accumulation of assets over the life cycle and the increase in the variance of labour productivity with age is important for the increase in the transitions from employment into non-employment late in the life cycle.

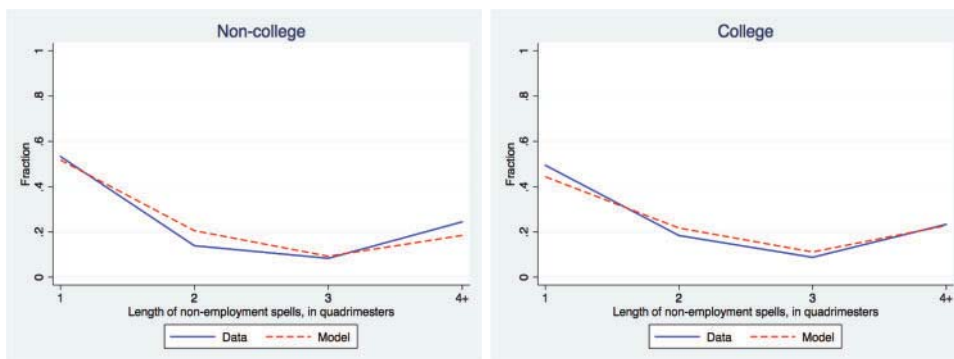


FIGURE 8
Non-employment spells, 1990 SIPP and model.

Non-employment spells. The length of non-employment spells provides a measure of the “volatility” of labour supply decisions along the extensive margin. If people that quit jobs go back to work the next period, then labour supply decisions along the extensive margin are quite “volatile.” We thus think it is important that the baseline economy is consistent with evidence on entry into non-employment as well as with the duration of non-employment spells. An important achievement of our theory is that it is consistent with evidence on both the incidence and the volatility of non-employment. Figure 8 shows that the model matches fairly well, qualitatively and quantitatively, the statistics on the length of non-employment spells documented on SIPP data. Non-employment spells are divided in four groups depending on their length: 1, 2, 3, or 4+ quadrimesters. For both education groups, in the SIPP and model data, more than 50% of all non-employment spells last one period and the fraction of non-employment spells lasting four or more periods is about 20%.

4.3.4. Moments not targeted in the calibration. The model also performs very well in dimensions not directly targeted in the calibration procedure.

Hours worked. The calibration targeted the employment rate for four age groups and the mean hours worked for employed prime-age males (35–50 years old). Figure 9 shows that the model matches the age profile of average hours worked in the population (including individuals with zero hours worked) quite well. Figure 10 shows that the overall distribution of hours worked in a quadrimester in the baseline model economy is similar to the distribution observed in the 1990 SIPP. In particular, the distribution is bimodal, the mass of individuals working zero hours increases with age, and the mass of individuals working full time (between 600 and 900 hours in a quadrimester) decreases with age.

Unemployment–employment (UE) flow rates. We now compare the UE flow rates in the model with those in the data. This comparison is interesting because our calibration of the job finding rate p did not directly target data on UE transitions. As discussed in Section 4.2.1, our calibration minimized a loss function that used information in the duration of non-employment spells. Alternatively, p can be identified with data on the job finding rate of unemployed individual. We argue, however, that this alternative approach will not change our results. We use the 1990 SIPP data set to compute the *monthly* flow rates from unemployment into employment (UE). For prime age males, we obtain a flow rate of 0.243 for non-college and 0.199 for college individuals.

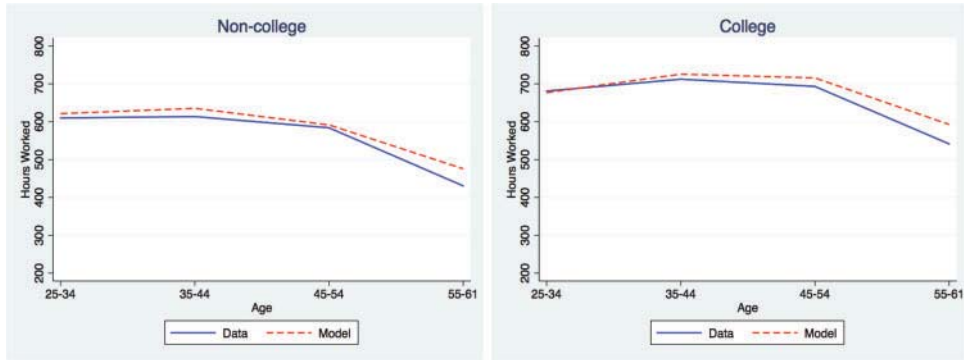


FIGURE 9
Mean hours, 1990 SIPP and model.

These monthly UE flow rates are consistent with evidence from other studies.²⁰ A 0.243 monthly UE flow rate implies that an individual that is unemployed in a given month will become employed within the next four months with probability 0.67.²¹ The quadrimesterly UE flow rate for non-college individuals in the model is 0.68. Similarly, the quadrimesterly UE flow rate for college individuals from the 1990 SIPP of 0.58 is basically the same as the one in the model. Therefore, the benchmark calibration is consistent with the UE flow rates observed in the 1990 SIPP data. In light of these results, we conclude that the findings in our article essentially will not change if we pin down p by targeting UE flow rates.

Non-employment spells by age. The calibration targeted statistics on the duration distribution of *all* non-employment spells. Figure 11 shows that the model captures quite well how the duration distribution of non-employment spells varies over the life cycle in the SIPP data. When individuals are young, both for non-college and college individuals, about 60% of the non-employment spells last for only one period. As individuals age, this fraction decreases and is about 40% for the age group 55–61. On the other hand, for both educational categories, the fraction of non-employment spells that are of length equal to four or more periods grows over the life cycle reaching a value close to 40% for the age group 55–61. The model matches these facts closely.

Cross-sectional heterogeneity in hours and lifetime labour supply by age. Consistent with the SIPP data, the model predicts that the coefficient of variation in hours of work is large, grows with age, and at the end of the life cycle reaches a value above 0.6 for non-college and of about 0.4 for college individuals (Figure 12). Recall that our calibration procedure implies that the variance of taste for leisure is constant over the life cycle. Hence, the increase in the variance of working hours with age is due to an increase in the variances of labour productivity and wealth over the life cycle. For all age groups, both in the model and in the data, heterogeneity in labour supply is larger for non-college than college individuals.

The calibration targeted the coefficient of variation of lifetime hours for the age group 35–44. The baseline economy matches remarkably closely how inequality in lifetime labour supply (the

20. Menzio *et al.* (2016), using the 1996 SIPP data set, report a monthly UE flow rate for high school males of 0.25; Choi *et al.* (2015), using data from the Current Population Survey (CPS), report a monthly flow rate for males in the range of 0.23–0.27; Krusell *et al.* (2010) use an UE flow rate of 0.25 in their analysis.

21. If the monthly UE transition rate is p_m , the probability that an unemployed individual *does not* transit into employment during a four-month period is $(1 - p_m)^4$. Conversely, the transition rate from unemployment into employment (UE transition rate) during a four-month period is then $1 - (1 - p_m)^4$.

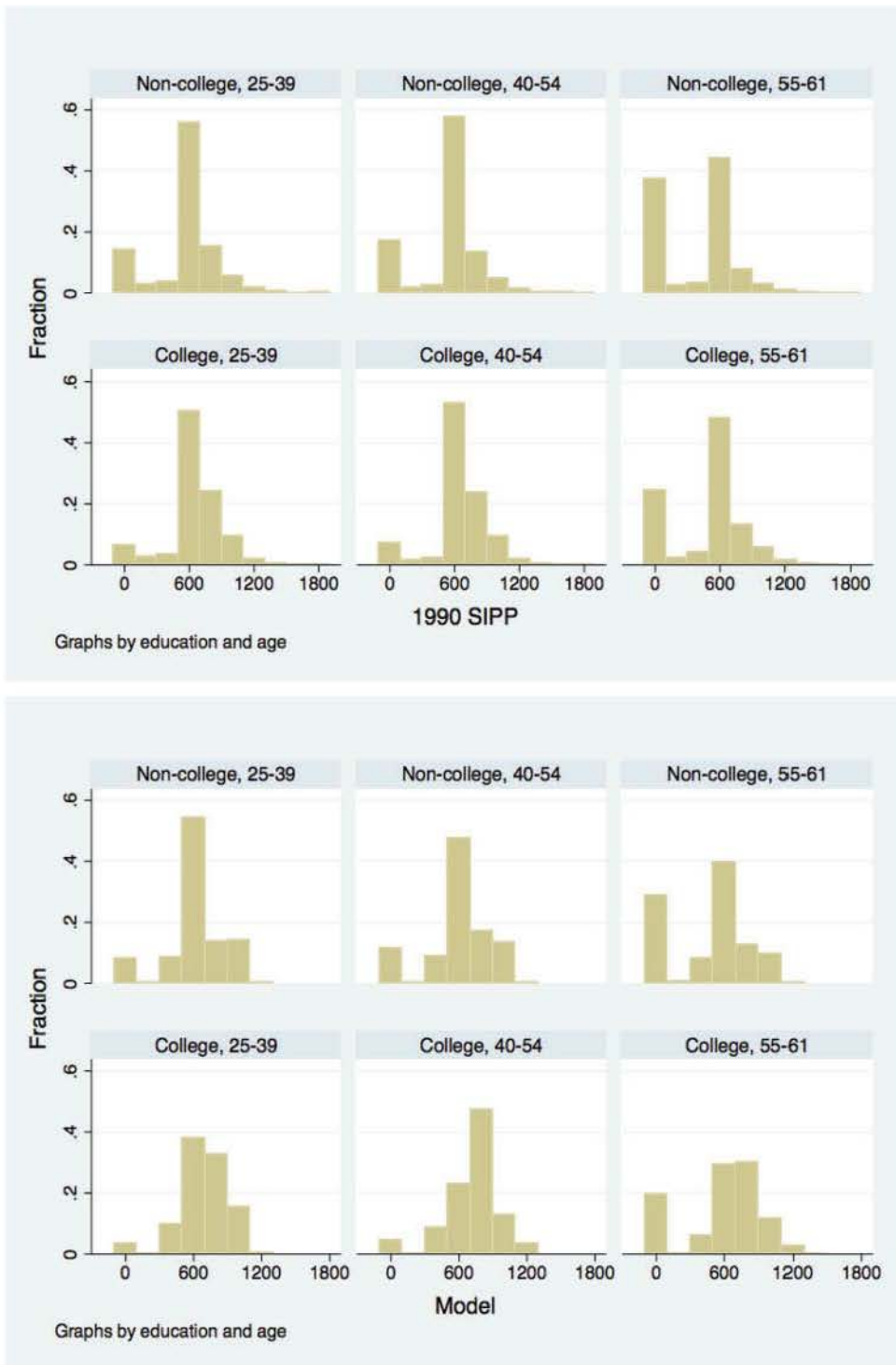


FIGURE 10
Histograms of hours worked in a quadrimester, 1990 SIPP and benchmark model.

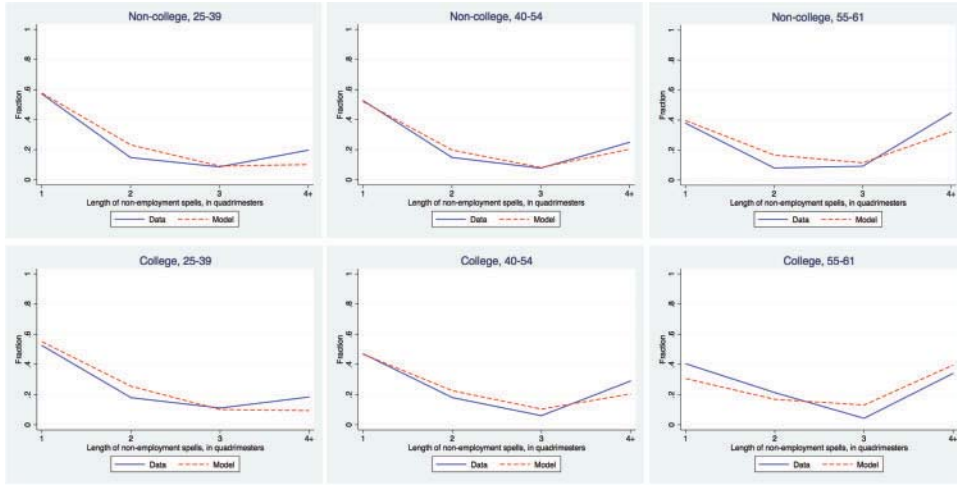


FIGURE 11
Non-employment spells, 1990 SIPP and model, by age.

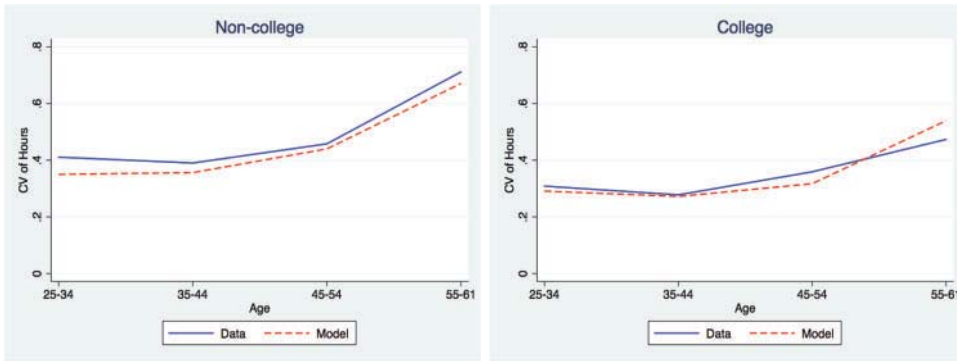


FIGURE 12
Coefficient of variation in hours, PSID and model.

Notes: Let h_{ij} denote the annual hours worked for individual i who is j years old. The coefficient of variation in hours across all individuals of age j is $CV^j(h_{ij})$. The figure reports the average $CV^j(h_{ij})$ over a 10-year period (e.g. ages 35–44), $E_j[CV^j(h_{ij})]$.

coefficient of variation over the sum of hours worked during a 10-year period) increases over the life cycle and that the increase in the inequality of lifetime hours over the life cycle is larger for non-college than college individuals. Indeed, as seen on Figure 13, quantitatively the model matches the age-profile of lifetime hours inequality very well for both education groups. Since the variance of preference shocks is constant over the life cycle, the increase in lifetime-hours inequality is explained by two mechanisms embedded in the model: life-cycle behaviour and incomplete markets. Individuals build precautionary savings early in the life cycle to ensure against idiosyncratic risk. As they age and accumulate assets, the heterogeneity in asset holdings translate into more persistent differences in labour supply across individuals.

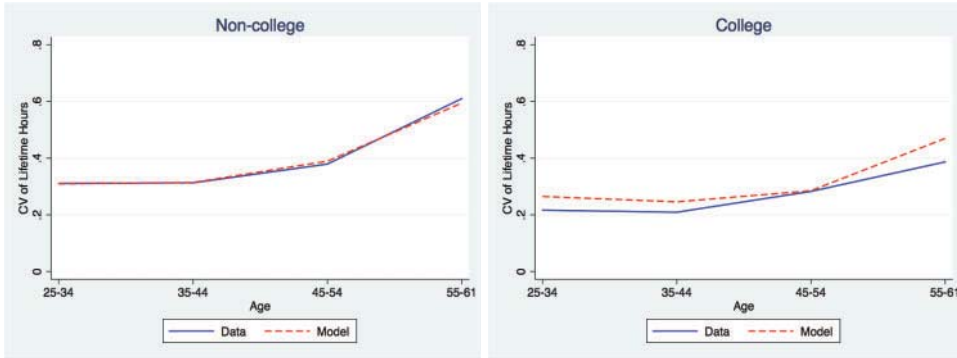


FIGURE 13

Coefficient of variation in lifetime hours, PSID and model.

Notes: For each 10-year period (*e.g.* between the ages of 35 and 44), let h_{ij} denote the annual hours worked for individual i who is j years old and let $\tilde{h}_i = (\sum_{j=35}^{44} h_{ij})$ denote the individual i 's lifetime hours during this period. The coefficient of variation in lifetime hours across all i individuals is $CV(\tilde{h}_i)$.

Consumption inequality over the life cycle. We have shown that the baseline economy is consistent with life-cycle data on U.S. inequality in wages, labour supply, and lifetime labour supply. Huggett *et al.* (2011) and Kaplan (2012a) advocate the view that a life-cycle theory of inequality and incomplete markets should be broadly consistent with evidence on consumption inequality over the life cycle. It is therefore interesting to compare the baseline economy's implications for the rise in consumption dispersion over the life cycle with the patterns found in U.S. data. The model economy implies a variance of log consumption of 0.17 at age 25 and of 0.30 at age 60. Hence, consumption inequality rises by 13 log points between the ages of 25 and 60. These implications are consistent with the evidence on consumption inequality reviewed in Huggett *et al.* (2011). In particular, Aguiar and Hurst (2013) document that the increase in the variance of consumption is about 12 log points when consumption is measured as total non-durable expenditures, with an initial value at age 25 of 0.15. The model's implications are remarkably close to these estimates. The Gini index of consumption in the baseline economy is 0.28, which is again close to the value of 0.26 reported by Krueger and Perri (2006) for the US economy during the 1990–2000 period.

5. AGGREGATE LABOUR SUPPLY RESPONSES

The analysis so far has shown that the baseline economy is able to match well facts on male labour supply over the life cycle along the extensive and intensive margins at the individual level. Therefore, the model economy is an appropriate tool for studying aggregate labour supply responses to changes in the economic environment. The model is used to explicitly aggregate up each individual's response, as well as to analyse how responses vary across subgroups in the population. Moreover, in a series of experiments we shut down one by one various modelling assumptions in order to evaluate their importance for understanding aggregate labour supply responses.

5.1. Response to a temporary wage change

We start by simulating in the baseline economy the aggregate labour supply response to a one-period (quadrimester) unanticipated wage change of 2%. In principle, one may think that the

TABLE 4
Aggregate labour supply elasticity to a temporary wage change

Elasticities	All			Non-college			College		
	Total	Intensive	Extensive	Total	Intensive	Extensive	Total	Intensive	Extensive
1. Baseline economy	1.75	0.67	1.08	1.96	0.67	1.29	1.35	0.66	0.69
2. Low productivity risk	1.42	0.69	0.73	1.60	0.69	0.91	1.09	0.68	0.41
3. No labour market frictions	1.85	0.68	1.17	2.04	0.69	1.35	1.49	0.66	0.83
4. No credit	1.51	0.67	0.84	1.64	0.67	0.97	1.27	0.67	0.60
5. Linear earnings + No Fixed Cost	0.85	0.85	0.0	0.86	0.86	0.00	0.81	0.81	0.00
6. No fixed cost	1.27	0.70	0.57	1.39	0.71	0.68	1.04	0.69	0.35
7. Linear earnings	0.85	0.84	0.01	0.86	0.85	0.01	0.82	0.81	0.01
8. No preference heterogeneity	1.29	0.71	0.58	1.37	0.71	0.66	1.13	0.70	0.43

wealth effect of such a change is negligible so that the change in the aggregate labour supply provides an estimate of the Frisch elasticity of aggregate labour supply. However, this reasoning is not firmly grounded in economic theory as with incomplete markets and borrowing constraints the Frisch elasticity of labour supply is an ill-defined concept.²² First, with borrowing constraints the Frisch elasticity does not exist because agents cannot hold the marginal utility of consumption constant as wages fluctuate. Secondly, with heterogeneity and incomplete markets a change in the wage rate will affect the marginal utility of wealth differently for different types of people. Hence, our results in this section should be viewed as describing aggregate labour supply responses to a temporary wage change rather than the Frisch elasticity of labour supply.^{23,24}

The aggregate elasticity of labour supply to a temporary wage change can be decomposed in terms of the elasticities along the intensive and extensive margins. The change in the average hours of work among employed individuals is used to compute the elasticity along the intensive margin. The elasticity along the extensive margin is given by the elasticity of the employment rate.²⁵ The results from these computations are reported in Table 4. We find that the elasticity of aggregate labour supply with respect to a temporary wage change in the baseline economy is 1.75. Restricting attention to labour supply changes along the intensive margin decreases the response from 1.75 to 0.67. Hence, the extensive margin accounts for about 62% of the aggregate labour supply response to a temporary wage change.

The elasticity of labour supply to a temporary wage change varies importantly across education and age groups. It is higher for high school than for college educated individuals (1.96 versus 1.35, respectively). Labour supply responses have a strong life-cycle pattern: the age profile of the elasticity is U shaped for both education categories. For high school individuals, the elasticity of labour supply takes a value of 2.01 for the 25–34 age group and decreases to 1.62 for the 35–44 age group. Afterwards, it increases to 1.90 for the 45–54 age group and to 2.74 for the 55–61 age group. The variation of the elasticity of labour supply for college individuals over the life cycle exhibits a similar U-shaped pattern. Such a pattern is reminiscent of the fact that over the business

22. See Keane (2011) for a discussion.

23. We thank an anonymous referee for pointing this out.

24. We note that in our baseline economy about 93.5% of all individuals experience a less than 0.2% change in their consumption in response to a 2% wage change implying that the variation caused by a wage change in the marginal utility of wealth across agents in the baseline economy is small.

25. Aggregate hours of work H can be expressed as $H = e \times h$, where e denotes the employment rate and h mean working hours. Taking logs and differentiating with respect to the log wage gives an expression for the aggregate elasticity in terms of the extensive and intensive margin elasticities: $\eta_H = \eta_e + \eta_h$.

TABLE 5
Extensive margin elasticity to a temporary wage change, by age and education

Panel A: non-college	25–34	35–44	45–54	55–61
1. Baseline economy	1.33	0.95	1.23	2.08
2. Low productivity risk	1.27	0.89	0.82	0.46
3. No labour market frictions	1.09	0.94	1.45	2.68
4. No credit	0.52	0.72	1.21	2.14
5. Linear earnings + no fixed cost	0.00	0.00	0.00	0.00
6. No fixed cost	0.67	0.50	0.75	0.97
7. Linear earnings	0.00	0.00	0.00	0.05
8. No preference heterogeneity	0.30	0.34	0.62	2.28
Panel B: College	25–34	35–44	45–54	55–61
1. Baseline economy	0.64	0.45	0.58	1.42
2. Low productivity risk	0.60	0.40	0.27	0.39
3. No labour market frictions	0.47	0.56	0.72	2.09
4. No credit	0.31	0.42	0.59	1.46
5. Linear earnings + no fixed cost	0.00	0.00	0.00	0.00
6. No fixed cost	0.28	0.29	0.34	0.60
7. Linear earnings	0.00	0.00	0.00	0.04
8. No preference heterogeneity	0.27	0.26	0.31	1.18

cycle hours of work and employment fluctuate much more for young and old individuals than for the middle-aged, as discussed in Gomme *et al.* (2005) and Jaimovich and Siu (2009).

The variation of labour supply responses across education and age groups in the baseline economy is almost all due to changes along the extensive margin. To fix ideas, let us focus on high school individuals. The intensive margin elasticity is almost flat over the life cycle: It decreases monotonically from 0.69 at age 25–34 to 0.66 at age 55–61. On the contrary, the extensive margin elasticity has a U-shape over the life cycle: It starts at 1.33 for the age group 25–34, decreases to 0.95 at age 35–44 and then rises to 1.23 and to 2.08 at ages 45–54 and 55–61, respectively (Table 5). Hence, the U-shaped age profile of the labour supply response to temporary wage change is thus driven by the variation of labour supply responses along the extensive margin. A similar conclusion applies to college-educated individuals. Furthermore, the extensive margin is crucial for understanding differences in the elasticity of labour supply across education groups. In short, the intensive margin is quantitatively important, but is not a factor in understanding the variation in the elasticity of aggregate labour supply to a temporary wage change across age and education groups. All the variation in labour supply responses across these groups is due to the extensive margin.

5.1.1. Evaluating the importance of modelling assumptions for the response to a temporary wage change. The baseline economy models (1) productivity risk, (2) labour market risk, (3) credit markets, (4) fixed cost of work, (5) non-linear earnings, and (6) preference heterogeneity. We now conduct various experiments to isolate the importance of each of these modelling assumptions for the aggregate labour supply response to a temporary wage change. In each of these experiments, we shut down one model assumption and keep the rest of the parameters of the baseline economy fixed. We then evaluate the labour supply change to a one period unanticipated wage change of 2% in each of the modified model economies. The results from these experiments are summarized below.

Productivity risk. To isolate the role of labour productivity risk, we assume that the standard deviation of the innovation to the autoregressive process on labour productivity is 100 times

smaller. As a result, we obtain an economy in which individuals essentially do not face labour productivity risk (apart from the initial fixed effect). The rest of the parameters are kept equal to the ones in the baseline economy. Table 4, row 2, reports the elasticity of labour supply to a temporary wage change for the modified-baseline economy with no labour productivity risk. The elimination of productivity risk decreases the aggregate elasticity of labour supply from a value of 1.75 in the baseline economy to 1.42. This decrease is entirely due to a lower labour supply response at the extensive margin (from 1.08 to 0.73) and is mostly due to old individuals. When productivity risk is shut down, for both education groups, the elasticity along the extensive margin for individuals aged 55–61 decreases sharply: It decreases from 2.08 to 0.46 for non-college and from 1.42 to 0.39 for college individuals (Table 5). This result is explained through the effect of productivity risk on life-cycle savings in the baseline economy. Since individuals have stronger incentives to save in order to self-insure against productivity risk, old individuals in the baseline economy are richer than in the economy with no risk and their employment decisions become more responsive to temporary wage changes.

Labour market risk. To assess the importance of labour market risk, we consider an economy in which the probability of finding a job is set to one ($p = 1$), keeping all other parameters from the baseline economy constant. The results are reported in row 3 of Table 4. When labour market risk is shut down the elasticity of aggregate labour supply increases from 1.75 in the baseline economy to 1.85. Interestingly, this result differs from the previous experiment, where we found that shutting down labour productivity risk led to a *decrease* in the aggregate labour supply response. In understanding why labour market risk and productivity risk have opposite effects on labour supply decisions, it is important to keep in mind that labour market risk affects labour supply decisions through two channels. First, it encourages precautionary savings which, as discussed before, makes employment decisions of old individuals much more responsive to temporary wage changes. Now, the strength of this effect is much weaker in the case of labour market risk than productivity risk because in our baseline model economy labour market shocks (job finding rates) are not persistent and thereby have a weaker effect on precautionary savings. Secondly, by making it difficult to find a job, labour market risk makes it more costly for individuals to take periods off work and the extensive margin becomes less responsive to temporary wage changes. We find that the second effect quantitatively prevails, especially late in the life cycle. As a result, as indicated in Table 5, the labour supply elasticity of high school individuals aged 55–61 along the extensive margin *increases* from 2.08 to 2.68 when labour market risk is shut down (recall that this statistic *decreases* to 0.46 in the experiment eliminating labour productivity risk). For college individuals, this statistic *increases* from 1.42 in the baseline economy to 2.09 when labour market risk is eliminated (while it *decreases* to 0.39 in the absence of productivity risk).

Credit markets. In the baseline economy, individuals can borrow up to about 40% of average annual earnings in the economy. Shutting down credit markets, decreases the elasticity of aggregate labour supply from 1.75 in the baseline economy to 1.51 (see row 4 in Table 4). The impact of credit markets on labour supply responses is almost all due to changes along the extensive margin. These effects vary over the life cycle in important ways. Focusing on high school individuals, the elimination of credit markets *decreases* the labour supply response along the extensive margin of individuals aged 25–34 from 1.33 to 0.52, while it *increases* the extensive margin elasticity of individuals aged 55–61 from 2.08 to 2.14 (Table 5). Credit markets also have opposite effects on the labour supply responses of young and old individuals with college education. Since young individuals hold few assets, the absence of credit makes it more costly for young individuals to take periods off work, diminishing their labour supply responses along the extensive margin. On the other hand, the absence of credit encourages individuals to

build precautionary savings making their employment decisions more responsive to temporary wage shocks when old.

Non-linear earnings, fixed costs of work, and extensive margin responses. We now consider labour supply responses in three different economies. First, we consider an economy with no fixed costs of work ($\bar{F}=0$) and linear earnings ($\theta=1$). Since in this economy the extensive margin is essentially inoperative, the elasticity of aggregate labour supply decreases from 1.75 to 0.84 relative to the baseline economy (see row 5 in Table 4). The decrease in the aggregate labour supply response is all explained by a lower response along the extensive margin (from 1.08 to 0).²⁶ Hence, in an economy with no fixed costs of work and linear earnings the elasticity of labour supply along the extensive margin is zero. To evaluate separately the role of the fixed cost of work and non-linear earnings for the observed extensive margin responses, we evaluate employment responses when only one of these two model features is active. If we add the calibrated fixed cost of work to the economy with linear earnings, the elasticity along the extensive margin is essentially zero (see row 7 on Table 4), suggesting that the calibrated fixed cost of work is too small for the extensive margin to respond to a small temporary wage change.²⁷ If, instead, we add non-linear earnings to an economy with no fixed costs of work the extensive margin elasticity is 0.57, which is a substantial response but still half of the overall response of 1.08 (compare rows 1 and 6 in Table 4). The last observations underscores that the calibrated fixed costs of work are important for employment responses in the presence of non-linear earnings.

Preference heterogeneity. To evaluate how heterogeneity in preferences affects labour supply responses, we consider an economy in which the variance of the preference shocks is set to zero. When preference heterogeneity is shut down in the baseline economy, the elasticity of aggregate labour supply to a temporary wage change decreases from 1.75 to 1.29 (compare rows 1 and 8 in Table 4)). We find that eliminating preference heterogeneity reduces the aggregate employment elasticity by about a half (from 1.08 to 0.58), with this reduction affecting all education and age groups (Table 5).²⁸ These effects are particularly large for individuals aged 25–34: the extensive margin elasticity falls by more than a fourth for high school individuals (from 1.33 to 0.30) and by more than a half for college individuals (from 0.64 to 0.27). Preference heterogeneity leads to a higher employment elasticity because it allows for “lazy” types who value leisure strongly. These individuals are much less likely to work than “non-lazy” individuals and, conditional on working, they are much more likely to enter a non-employment spell. As a result, “lazy” types exhibit a high employment elasticity to temporary wage changes.

In summary, the impact of modelling assumptions on labour supply responses is almost entirely driven by their effects on decisions along the extensive margin. When we simulate employment responses to a small wage change in an economy with linear earnings, we find that the calibrated fixed costs of work are sufficiently small that, on their own, they do not affect employment decisions along the extensive margin. However, fixed costs of work matter importantly in the presence of non-linear earnings, accounting for half of the employment response in our baseline

26. The intensive margin elasticity is higher in the economy with an inactive extensive margin due to the fact that working hours are lower in this economy. This can be understood by recalling that $\eta^h = \frac{1-h}{h} \eta^l$, where η^h and η^l denote the elasticity of hours and leisure, and h denotes hours of work. Since the elasticity of leisure is fixed at $1/\sigma$, a decrease in h mechanically increases the elasticity of hours.

27. We will further discuss this issue in Section 5.4.2.

28. The only exception is for high school individuals aged 55–61, whose employment elasticity increases from 2.08 to 2.28 without preference heterogeneity.

economy. This result underscores that there is a complementarity between fixed costs of work and non-linear earnings that enhances the aggregate labour supply response to temporary wages.

The extensive margin is also crucial for understanding how the labour supply responses to temporary wage change varies across age and education groups. The age profile of the labour supply elasticity in the baseline economy is U shaped. Modelling credit (or initial assets), preference heterogeneity, and the extensive margin are important for generating the decline in the elasticity early in the life cycle. If any of these features are shut down in the baseline economy, the elasticity of labour supply does not decrease with age when individuals are young. Modelling productivity risk and the extensive margin (non-linear earnings) are important for accounting for the increase in the labour supply response to a temporary wage change late in the life cycle. Preference heterogeneity increases aggregate labour supply responses, and this effect is strongest for young individuals with high school education. Finally, labour market frictions have a large negative effect on the labour supply response of old individuals.

Our model abstracts from human capital accumulation. It is interesting to point out that in a model with human capital accumulation, as in Imai and Keane (2004) and Keane (2011), the labour supply response to a temporary wage change increases over the life cycle. Young individuals do not respond much to temporary wage shocks because the opportunity cost of work is not only the current wage but also the return to human capital accumulation. As individuals age, the returns to human capital accumulation decline and individuals respond more to temporary wage changes. Our model focuses on different, but complementary, mechanisms that have similar implications as the ones in Imai and Keane (2004). For instance, the need to build precautionary savings makes young individuals less responsive to temporary wage changes while as they age and accumulate assets they become more responsive.

5.1.2. Discussion. Since non-linear earnings is a relatively unexplored feature, we find it interesting to evaluate its role by calibrating a version of our baseline economy with linear earnings (note that in the previous section linear earnings were introduced but without re-calibrating the model economy). In the Online Appendix C, we compare the re-calibrated economy with linear earnings with the baseline economy, both in terms of their fit of the micro-facts as well as in terms of their implications for aggregate labour supply responses to a temporary shock.²⁹ In particular, we compare the response to a one period (quadrimester) wage change of 2% across the two calibrated model economies. We find that the economy with linear earnings has a much lower aggregate response than the baseline economy, with an aggregate labour supply elasticity of 1.05 versus 1.75 in the baseline economy. The lower response in the re-calibrated economy with linear earnings is explained by the extensive margin. In the re-calibrated economy with linear earnings the elasticity of aggregate labour supply along the extensive margin is less than a half the one in the baseline economy (0.44 versus 1.08). We draw two key lessons: first, fixed costs of work may allow for significant (bounded away from zero) extensive margin responses if they are large enough. Secondly, non-linear earnings *amplify* the effect of fixed costs of work on

29. We find that the economy with linear earnings implies a worse fit of the calibration targets. The most striking difference is that the economy with linear earnings requires fixed costs of work about four times higher than in the baseline economy: for instance, focusing on high school individuals, the economy with linear earnings requires a fixed cost of work equivalent to a consumption loss of 36%, instead of the 8.5% value in the baseline economy. In the Online Appendix C, we discuss evidence on work-related expenditures from Aguiar and Hurst (2013) suggesting that the fixed costs of work in the baseline economy are empirically more plausible. In the Online Appendix C, we also perform a sensitivity analysis by re-calibrating the baseline economy under the assumption that the parameter σ , determining the intertemporal substitution of leisure, is equal to three (rather than the value of two assumed in the calibration of the baseline economy). We also compare labour supply responses to permanent (compensated) wage (tax) changes.

extensive margin responses to a temporary wage change. The baseline economy (with non-linear earnings) has an extensive margin response to a 2% wage increase that is 2.5 times higher (*e.g.* the ratio between 1.08 and 0.44 is about 2.5) than the re-calibrated economy with linear earnings even though fixed cost of work are about a factor of four smaller in the former economy.

Although restricting the analysis to 25–61 old individuals is not unusual, it does warrant a discussion. We exclude individuals younger than 25 in order to make sure that they have already completed their schooling. Individuals older than 61 years are excluded since at age 62 individuals can start collecting social security payments. In many ways, our model and its computation are greatly simplified by focusing on the 25–61 age group. On the one hand, incorporating in the analysis individuals younger than 25 would require jointly modelling the education and labour supply decisions of young individuals, as in Keane and Wolpin (2001), for instance. These authors also point out that a careful modelling of schooling decisions requires explicitly accounting for heterogeneity in preferences for schooling, tuition costs, and parental transfers. Moreover, Kaplan (2012b) argues that the fact that young people have the option to live with their parents affects their consumption, saving, and labour supply decisions. On the other hand, incorporating individuals older than 61 would require modelling in much greater detail the retirement decisions and social security rules. Incorporating all these important additional aspects in the model would add substantial complexity to the model and its computation, and as a result we opted for concentrating on the 25–61 age group. To the extent that individuals younger than 25 are more likely to be indifferent about working at the offered wage due to the fact that they have the option to live with their parents, as argued in Kaplan (2012b), the labour supply response (especially at the extensive margin) would be higher than in our benchmark economy. Similarly, we expect the labour supply responses of older individuals, collecting social security payments, to be quantitatively more pronounced.

The empirical literature on labour supply provides a broad set of estimates for the intertemporal elasticity of labour supply along the intensive margin. Keane (2011) surveys 21 of the best known studies that estimate the (Frisch) elasticity of labour supply at the intensive margin and reports a mean value from the surveyed studies of 0.83. Chetty *et al.* (2011a), in their meta study of 25 papers, report a mean value of 0.54 for the (intensive margin) Frisch elasticity of labour supply. Our theory implies a value of 0.67, which is consistent with the findings in this literature. Furthermore, we stress that the intertemporal elasticity in our model economy is computed from an ideal experiment for which there is no clear counterpart in the data: We simulate a one period small (purely) unanticipated wage change and compute aggregate labour supply responses to measure the intertemporal elasticity of labour supply. If we were to measure the intertemporal response using conventional empirical methods on the model simulated household-level data (aggregated to the annual level), assuming no measurement error in hours and wages whatsoever, we would have obtained a Frisch elasticity of labour supply in the range of 0.17 to 0.57, depending on the instruments³⁰ used in the log-hours-log-wage regression (empirical) analysis. Allowing for some empirically reasonable measurement error in hours produces estimates in the range of -0.79 to 0.55 , which is again consistent with the estimates in the empirical literature. The fact that the conventional empirical methods leads to estimates that are lower than the elasticity obtained with our ideal macro-experiment is consistent with the views discussed in Keane (2011) and Keane and Rogerson (2012).

The cross-country heterogeneity in labour supply among individuals of old age and social security provisions present us with the opportunity to test the predictions of our theory. In Erosa *et al.* (2012), we use a related framework with non-linear earnings to model in detail

30. We considered standard instruments from the literature (lagged log wage changes; or twice lagged log wage changes; or age, age squared, and twice lagged log wage; or cubic in age). See Imai and Keane (2004) and Keane (2011).

TABLE 6
Aggregate labour supply elasticity to a permanent-compensated wage change

	Total	Extensive	Intensive
Baseline economy	0.44	−0.11	0.55
Linear earnings	0.56	0.00	0.56
No preference heterogeneity	0.49	−0.07	0.56

the variation in the social security, disability insurance, and taxation institutions across European countries and the U.S. We find that the model economy accounts well for the observed cross-country differences in labour supply late in the life cycle, indicating that the intertemporal labour supply responses in our framework are plausible.

Our findings suggest that time aggregation is an important issue. The consensus in the literature, to some extent, has been that the extensive margin at the annual level is not quantitatively important for males. One contribution of our article is to show that individuals frequently use the extensive margin to adjust their labour supply within a year, even at old ages. These extensive-margin responses are not as prominent once we aggregate male labour supply to the annual level. Therefore, capturing these pattern is central to the analysis in this article, and we find that it is essential for understanding labour supply responses. Conceptually then, a four-month model period might also miss some higher-frequency extensive margin fluctuations: individuals who work only one or two months during a quadrimester do respond at the extensive margin, but at the quadrimester level this will be interpreted as intensive margin fluctuation in labour supply. In order to understand whether this would lead to a higher or smaller aggregate elasticity of labour supply, we would ideally have to formulate our theory with a smaller period—*e.g.* a month or a week. Unfortunately, this is infeasible at this moment. The computational cost would simply be enormous.³¹ Nonetheless, we believe that it is an important contribution of our article to show that there are important changes along the extensive margin within a year (even if just focusing on a quadrimester level and ignoring monthly or weekly fluctuations in labour supply). Our findings in the SIPP data show that the extensive margin is quite active at the quadrimester level. Relative to the state of the art, we think that we have made substantial progress by looking at a four month, instead of an annual, period.

5.2. *Response to a permanent compensated wage (tax) change*

In his survey of labour supply and taxes Keane (2011) argues that the Hicks elasticity measures the behavioural response that determines the deadweight losses of taxing labour income and thus is crucial for the optimal design of the tax system. We now evaluate the aggregate labour supply response to a permanent compensated wage (or equivalent tax change), which provides an estimate of the Hicks elasticity of labour supply. We simulate a permanent wage decrease of 10% (or a an equivalent tax increase) together with a lump sum transfer that effectively compensates the decrease in earnings due to the lower wage rate (higher tax rate). Hence, the wage (tax) change affects the first-order condition for working hours and earnings are only affected to the extent

31. Note that our model features many dimensions of heterogeneity (education, fixed effects on labour productivity, fixed effects on taste shocks, shocks to labour productivity and shocks to preferences, life cycle), incomplete markets, and non-convex labour supply decisions (the first-order conditions have two interior solutions). Moreover, the calibration requires pinning down a large number of parameters and involves a nested iterative procedure (the inner loop requires the minimization of a loss function) which is computationally quite costly.

that individuals change hours of work. To put it differently, the compensated wage (tax) change does not directly affects earnings but only indirectly through its impact on hours of work.

We find that the Hicks elasticity of aggregate labour supply in the baseline economy is 0.44.³² Interestingly, labour supply responses along the intensive and extensive margins have different signs. While the intensive margin elasticity is 0.55, the extensive margin elasticity is -0.11 . Hours along the intensive margin decrease because of the substitution effect associated with the lower wage rate. On the other hand, the decrease in the (compensated) wage discourages savings over the life cycle which, in turn, leads to an important increase of the employment rate at old age. This effect accounts for the aggregate negative elasticity of labour supply along the extensive margin. In particular, the Hicks elasticity along the extensive margin for individuals aged 55–61 is -0.42 and -0.35 for non-college and college individuals. The employment elasticity for younger age groups is much smaller (less than one-fifth the value for the oldest age group).

Our findings suggest that modelling the extensive margin is important for an accurate assessment of aggregate labour supply responses to permanent wage (tax) changes. Neglecting the extensive margin response will overstate the aggregate labour supply response by one-fourth. The magnitude of this bias varies across ages and is largest for the oldest age group (by a factor of more than 3).

We also evaluated the Hicks elasticity in two separate experiments: In the first experiment, we shut down non-linear earnings (relative to the baseline economy), whereas in the second experiment we shut down preference heterogeneity. We find that in the case with linear earnings the aggregate Hicks elasticity is 0.56, which is higher than the 0.44 estimated for the baseline economy. The Hicks elasticity is now higher than in the baseline economy because the extensive margin elasticity is zero instead of negative. Shutting down preference heterogeneity also increases the aggregate Hicks elasticity from 0.44 in the baseline economy to 0.49. The aggregate Hicks elasticity is now slightly higher because the extensive margin elasticity decreases from -0.11 in the baseline economy to -0.07 in the absence of preference heterogeneity. Hence, unlike the findings for the response to a temporary wage change, preference heterogeneity does not matter much for the Hicks elasticity of labour supply.

Finally, unlike our findings, Rogerson and Wallenius (2009) (henceforth RW) find a positive extensive margin elasticity with respect to a permanent (compensated) wage (tax) change. Our model differs from RW in many dimensions. We find it important to highlight two key factors that account for the differences in results. First, we model the transfers that compensate workers for the tax increase (wage decrease) in a different way than RW. While RW compensate all individuals (regardless of whether they work or not) with a lump sum transfer that is equal to the amount of taxes collected at each age, in our computational experiment we give back to each worker a lump sum transfer equal to the actual increase in taxed paid (or, equivalently, a transfer that offsets the wealth effect of the permanent wage change). This implies that in our experiment individuals that do not work do not receive any transfer, while in the RW experiment non-working individuals do receive transfers. As a result, not surprisingly, a larger fraction of individuals choose not to work after a tax increase in RW than in our model. We believe that our compensation provides a more accurate estimate of the Hicks elasticity of labour supply (*e.g.* given the substantial amount of heterogeneity in our model economy, the RW transfer would have a large redistributive effect). Our baseline experiment implies a Hicks elasticity of labour supply along the extensive margin

32. This number is consistent with the evidence summarized in Chetty *et al.* (2011b) and in Keane (2011). Interestingly, however, Keane (2011) shows that in a model with human capital accumulation it is important to distinguish between the short-run and long-run effects of a compensated permanent tax change. Since a permanent increase in taxes decreases the incentives to accumulate human capital, the effects of such a tax reform on labour supply grow over time and are thus higher in the long run than in the short run.

of -0.10 . If we model a lump sum transfer that does not depend on whether individuals work or not³³, then the Hicks elasticity along the extensive margin rises to 0.10.

Secondly, differently from RW, our model economy features incomplete markets and borrowing constraints. If we shut down the idiosyncratic risk in the economy (and allow individuals to borrow up to 4 times their quadrimesterly income) we obtain that the extensive margin elasticity rises from 0.10 to 0.26. These findings point to the fact that the extensive margin response to a permanent tax (wage) increase (decrease) is lower with incomplete markets, such as our baseline economy. In the presence of incomplete markets, the transfer diminishes the incentives of individuals to build precautionary savings. The decrease in the life-cycle profile of asset holdings then encourages old individuals to work (*e.g.* their lower wealth discourages them from taking periods off work). This effect mitigates the negative employment response to the tax increase (wage decrease), and leads to the lower labour supply elasticity along the extensive margin under incomplete markets.

5.3. *The Icelandic tax experiment*

We test the predictions of our theory with evidence from a tax reform that took place in Iceland. In 1987 Iceland moved from a system under which taxes were paid on the previous year's income to a pay-as-you-earn system. The transition to the new tax system implied that income during 1987 was never taxed since the tax base in 1987 was income earned in 1986 and the tax base in 1988 was income earned in 1988. Chetty *et al.* (2011b) argued that this tax holiday induced an unanticipated temporary wage variation that provided a natural experiment for identifying Frisch elasticities. They simulated a one year unanticipated tax relief in the Rogerson and Wallenius (2009) framework and concluded that the RW model implied a labour supply response much larger (about a factor of five) than the estimated response in the data.

Our model economy is not calibrated to Iceland (and neither was the RW model). Nonetheless, the findings of Chetty *et al.* (2011b) raise the concern that our theory may be at odds with the evidence from the Icelandic Tax holiday. To evaluate this possibility, we simulate the Icelandic tax holiday in our baseline economy. The average tax rate in Iceland was 14.5% in 1986, 0% in 1987, and 8.0% in 1988.³⁴ In order to mimic the tax reform in Iceland, we simulate in our baseline economy a one year (three model periods) reduction in the tax rate of 14.5 percentage points, followed then by a permanent decrease of 6.5 percentage points in the average tax rate. Before proceeding to the results, we want to clarify the objective of this experiment. We emphasize that the Icelandic tax holiday does not provide evidence on Frisch elasticities of labour supply or the aggregate labour supply response to a transitory wage change. One reason is due to the fact that Iceland was undergoing substantial policy reforms in 1987, which were likely to affect future growth prospects and, in particular, affect the marginal utility of wealth of the agents in the economy. Further, as we have stated before, the 1987 tax holiday was accompanied by some permanent tax changes after 1987. The Icelandic tax holiday thus combines a temporary tax relief with a permanent decrease in taxes. Our goal then is to test if our theory is quantitatively consistent with the labour supply responses in the Icelandic data.

We find that the aggregate elasticity of labour supply implied by the Icelandic tax holiday experiment is 0.64. Remarkably, this elasticity result is not much different from the ones estimated

33. To minimize the redistributive effect of transfers, the transfers depend on the fixed characteristics of individuals—*e.g.* the fixed effects in labour productivity and the taste for leisure.

34. The tax change was unanticipated by households since the announcement of the policy change was made in late 1986, see Bianchi *et al.* (2001).

by Bianchi *et al.* (2001) in the Icelandic micro-data who found an aggregate labour supply elasticity of 0.84 for male workers.³⁵

It is instructive to compare the labour supply response in the Icelandic tax simulation with the one obtained when simulating a one quadrimester wage change of 2%. As we discussed earlier, simulating the effects of a one period (quadrimester) wage increase of 2% delivers an aggregate elasticity of labour supply equal to 1.75. (Table 4). The labour supply response is larger than the one obtained in the Icelandic experiment (0.64) for three reasons. First, the change in wages lasts only for one period (quadrimester) rather than three model periods (one year) and the scope for intertemporal substitution is higher in the case of a one period wage change. Moving from a one quadrimester to a one year wage change of 2% decreases the labour supply elasticity from 1.75 to 1.54. Secondly, the labour supply response in our model is non-linear in the size of the wage change: the labour supply response to a one-year wage change drops from 1.54 to 0.83 when the size of the wage change increases from 2% to 14.5% because the wealth effect is non-negligible for large wage changes. Thirdly, the Icelandic tax experiment combines a one-year temporary change with a permanent wage change, which further decreases the labour supply elasticity from 0.83 to 0.64.

6. CONCLUSION

We develop a life-cycle model with heterogeneous agents making labour supply decisions, both along the intensive and extensive margins, at subannual periods. The model is consistent with evidence from the SIPP that male individuals frequently use the extensive margin to adjust their labour supply within a year and with the large amount of heterogeneity in lifetime labour supply that we document using PSID data. As a result, the model provides an aggregation theory of individual labour supply that is firmly grounded on individual-level micro-evidence and is used to study aggregate labour supply responses. We find that the aggregate labour supply elasticity to a transitory wage shock is 1.75, with the extensive margin accounting for 62% of the response. Furthermore, we find that the aggregate labour supply elasticity to a permanent-compensated wage change is 0.44.

Our results provide insights for both business cycle and policy analysis. Modelling an operative extensive margin (within a year) is crucial for understanding labour supply responses of men. In the absence of an extensive margin, the response to a small one period wage shock (intertemporal elasticity of aggregate labour supply) would be underestimated by more than a half (0.67 instead of 1.75). The extensive margin response in our baseline economy is enhanced because of the interaction between fixed costs of work and non-linear earnings. In the absence of any of these two features, extensive margin responses are cut by more than a half. Moreover, we find that the age profile of the labour supply responses to temporary wage changes in our model economy is U shaped, which suggests that the ingredients in our model (such as preference heterogeneity and extensive margin) may be important in accounting for the fact that over the business cycle hours of work and employment fluctuate much more for young and old individuals than for the middle aged. Hence, in future work it would be interesting to embed our model in a business cycle framework.

35. This value comes from adding the extensive and intensive margin elasticities. They report an extensive margin elasticity of 0.58 in Table 4 in Bianchi *et al.* (2001). The intensive margin elasticity of 0.26 is obtained using data from Table 6. In our simulations, the response along the extensive margin is 0.23, which is substantially smaller than the 0.58 value reported by Bianchi *et al.* (2001). However, we should note that these authors used data on number of weeks worked to measure the extensive margin response. Since our model period is a quadrimester, our higher response along the intensive margin is probably capturing an extensive margin response in the data studied by Bianchi *et al.* (2001). We thus think that it is more appropriate to focus our analysis on aggregate labour supply responses.

We believe that our framework contains crucial ingredients for studying how tax and transfers institutions across countries affect aggregate labour supply. When the extensive margin is active, the effects of taxes on labour supply are non-linear in the size of the tax rates and these non-linearities may vary across individuals with different characteristics, such as age, education, and taste for leisure. Responses along the extensive margin have the potential of being large among those groups of individuals that have a low employment rate, but nonetheless they could be small if the tax changes are small and individuals' (after-tax) wages are far from their reservation values. Moreover, labour supply responses depend importantly on how the tax revenue is rebated back to consumers. Hence, to accurately assess the impact of taxes on labour supply, we believe that it is important to incorporate an active extensive margin, preference heterogeneity, and model the institutions determining how transfers are made to individuals. These features should also matter for the design of optimal tax policies and income support policies. These issues are left for future research.

Our theory emphasizes the importance of variations in the extensive margin, at the quadrimester level (four-month period), for understanding male labour supply. A four-month model period might also miss some higher-frequency extensive margin fluctuations: individuals who work only one or two months during a quadrimester do respond at the extensive margin, but at the quadrimester level these responses will be interpreted as intensive margin fluctuations in labour supply. In order to understand whether this would lead to a higher or smaller aggregate elasticity of labour supply, we would ideally have to formulate our theory with a smaller period—*e.g.* a month or a week. Unfortunately, the computational costs are too high at this moment, but we think that we have made an important contribution by assessing the importance of changes along the extensive margin within a year.

The insights in our article should also be quite relevant for the study of female labour supply, and extending the model to include female labour supply will enhance our understanding of aggregate labour supply responses. Finally, incorporating a household structure in the analysis in which both spouses make joint labour supply decisions would enhance our understanding of the determinants of labour supply responses and would allow for the realistic inclusion of various government policies and transfers such as Food Stamps and the Earned Income Tax Credit.

Acknowledgments. This version: January 15, 2016. We thank the Editor and three anonymous referees for helpful comments, as well as participants in numerous seminars and conferences. Andrés Erosa and Luisa Fuster acknowledge financial support from the Ministerio de Economía y Competitividad of Spain (grants ECO2012-31653 and MDM 2014-0431) and Comunidad de Madrid, MadEco-CM (S2015/HUM-3444). Luisa Fuster also acknowledges financial support from the European Community under the Marie Curie International Reintegration grant PIRG03-GA-2008-231099. Gueorgui Kambourov has received funding from the Social Sciences and Humanities Research Council of Canada grants #410-2011-1051 and #435-2014-0815 and from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n. 324085. We thank Hui Xiong for excellent research assistance.

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online.

REFERENCES

- AARONSON, D. and FRENCH, E. (2004), "The Effect of Part Time Work on Wages: Evidence from Social Security Rules", *Journal of Labour Economics*, **22**, 329–252.
- AGUIAR, M. and HURST, E. (2013), "Deconstructing Life Cycle Expenditure", *Journal of Political Economy*, **121**, 437–492.
- BHATTACHARYA, J. and LAKDAWALLA, D. (2006), "Does Medicare Benefit the Poor?", *Journal of Public Economics*, **90**, 277–292.
- BIANCHI, M., GUDMUNDSSON, B. R. and ZOEGA, G. (2001), "Iceland's Natural Experiment in Supply-Side Economics", *American Economic Review*, **91**, 1564–1579.

- BLUNDELL, R., BOZIO, A. and LAROQUE, G. (2011), "Extensive and Intensive Margins of Labour Supply: Working Hours in the US, UK and France" (IFS Working Paper 01/11, Institute for Fiscal Studies).
- CHANG, Y. and KIM, S.-B. (2006), "From Individual to Aggregate Labour Supply: A Quantitative Analysis Based on a Heterogeneous Agent Macroeconomy", *International Economic Review*, **47**, 1–27.
- CHETTY, R., GUREN, A., MANOLI, D., *et al.* (2011a), "Are Micro and Macro Labour Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins", *American Economic Review Papers and Proceedings*, **101**, 471–475.
- CHETTY, R., GUREN, A., MANOLI, D. and WEBER, A. (2011b), "Does Indivisible Labour Explain the Difference Between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities" (Working Paper No. 16279, National Bureau of Economic Research).
- CHOI, S., JANIAK, A. and VILENNA-ROLDÁN, B. (2015), "Unemployment, Participation and Worker Flows Over the Life Cycle", *Economic Journal*, forthcoming.
- COGAN, J. (2001), "Fixed Costs and Labour Supply", *Econometrica*, **49**, 945–963.
- COOLEY, T. F., editor (1995), *Frontiers of Business Cycle Research* (Princeton: Princeton University Press).
- DOMEIJ, D. and FLODÉN, M. (2006), "The Labour-Supply Elasticity and Borrowing Constraints: Why Estimates are Biased", *Review of Economic Dynamics*, **9**, 242–262.
- DOMEIJ, D. and HEATHCOTE, J. (2004), "On The Distributional Effects of Reducing Capital Taxes", *International Economic Review*, **45**, 523–554.
- EROSA, A., FUSTER, L. and KAMBOUROV, G. (2012), "Labour Supply and Government Programs: A Cross-Country Analysis", *Journal of Monetary Economics*, **59**, 84–107.
- FERNÁNDEZ-VILLAVEDE, J. and KRUEGER, D. (2007), "Consumption Over the Life Cycle: Facts from the Consumer Expenditure Survey Data", *The Review of Economics and Statistics*, **89**, 552–565.
- FRENCH, E. (2005), "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behavior", *Review of Economic Studies*, **72**, 395–427.
- GOMME, P., ROGERSON, R., RUPERT, P., and WRIGHT, R. (2005), "The Business Cycle and the Life Cycle", *NBER Chapters in: NBER Macroeconomics Annual 2004*, vol. 19, in Gertler, M. and Rogoff, K., (eds), National Bureau of Economic Research, Inc, p. 415–592.
- GUSTMAN, A. L. and STEINMEIER, T. L. (1985), "The Effect of Partial Retirement on the Wage Profiles of Older Workers", *Industrial Relations*, **24**, 257–265.
- GUSTMAN, A. L. and STEINMEIER, T. L. (1986), "A Structural Retirement Model", *Econometrica*, **54**, 555–584.
- HANSEN, G. (1985), "Indivisible Labour and the Business Cycle", *Journal of Monetary Economics*, **16**, 309–327.
- HEATHCOTE, J., STORESLETTEN, K. and VIOLANTE, G. L. (2014), "Consumption and Labour Supply with Partial Insurance: An Analytical Framework", *American Economic Review*, **104**, 1–52.
- HORNSTEIN, A. and PRESCOTT, E. C. (1993), "The Firm and the Plant in General Equilibrium Theory", in Becker, R., Boldrin, M., Jones, R. and Thomson, W., (eds), *General Equilibrium, Growth, and Trade II: The Legacy of Lionel McKenzie* (San Diego: Academic Press).
- HUGGETT, M., VENTURA, G. and YARON, A. (2011), "Sources of Lifetime Inequality", *American Economic Review*, **101**, 2923–2954.
- IMAI, S. and KEANE, M. P. (2004), "Intertemporal Labour Supply and Human Capital Accumulation", *International Economic Review*, **45**, 601–641.
- JAIMOVICH, N. and SIU, H. (2009), "The Young, the Old, and the Restless: Demographics and Business Cycle Volatility", *American Economic Review*, **99**, 804–826.
- KAPLAN, G. (2012a), "Inequality and the Lifecycle", *Quantitative Economics*, **3**, 471–525.
- KAPLAN, G. (2012b), "Moving Back Home: Insurance Against Labour Market Risk", *Journal of Political Economy*, **120**, 446–512.
- KAPLAN, G. and VIOLANTE, G. L. (2010), "How Much Consumption Insurance Beyond Self-Insurance?", *American Economic Journal: Macroeconomics*, **2**, 53–87.
- KEANE, M. P. (2011), "Labour Supply and Taxes: A Survey", *Journal of Economic Literature*, **49**, 961–1075.
- KEANE, M. P. and ROGERSON, R. (2011), "Reconciling Micro and Macro Labour Supply Elasticities: A Structural Perspective" (Working Paper No. 17430, National Bureau of Economic Research).
- KEANE, M. P. and ROGERSON, R. (2012), "Micro and Macro Labour Supply Elasticities: A Reassessment of Conventional Wisdom", *Journal of Economic Literature*, **50**, 464–76.
- KEANE, M. P. and WOLPIN, K. I. (2001), "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment", *International Economic Review*, **42**, 1051–1103.
- KRUEGER, D. and PERRI, F. (2006), "Does Income Inequality Leads to Consumption Inequality? Evidence and Theory", *Review of Economic Studies*, **73**, 163–193.
- KRUSELL, P., MUKOYAMA, T., ROGERSON, R., and SAHIN, A. (2010), "Aggregate Labour Market Outcomes: The Roles of Choice and Chance", *Quantitative Economics*, **1**, 97–127.
- LOW, H. (2005), "Self-Insurance in a Life-Cycle Model of Labour Supply and Savings", *Review of Economic Dynamics*, **8**, 945–975.
- LOW, H., MEGHIR, C., and PISTAFERRI, L. (2010), "Wage Risk and Employment Risk Over the Life Cycle", *American Economic Review*, **100**, 1432–1467.
- McDANIEL, C. (2007), "Average Tax Rates on Consumption, Investment, Labour and Capital in the OECD 1950–2003" (Mimeo, ArizonaStateUniversity).
- MEGHIR, C. and PISTAFERRI, L. (2004), "Income Variance Dynamics and Heterogeneity", *Econometrica*, **72**, 1–32.

- MEGHIR, C. and PISTAFERRI, L. (2011), “Earnings, Consumption and Life Cycle Choices”, in Card, D. and Ashenfelter, O., (eds), *Handbook of Labour Economics*, 4B (Amsterdam: Elsevier).
- MENZIO, G., TELYUKOVA, I. A. and VISSCHERS, L. (2016), “Directed Search Over the Life Cycle”, *Review of Economic Dynamics*, forthcoming.
- MOFFITT, R. (1984), “The Estimation of a Joint Wage-Hours Labour Supply Model”, *Journal of Labour Economics*, 2(4), 550–566.
- OSUNA, V. and RÍOS-RULL, J.-V. (2003), “Implementing the 35 Hour Workweek by Means of Overtime Taxation”, *Review of Economic Dynamics*, 6, 179–206.
- PIJOAN-MAS, J. (2006), “Precautionary Savings or Working Longer Hours?”, *Review of Economic Dynamics*, 9, 326–352.
- PRESCOTT, E. (2004), “Why Do Americans Work Much More than Europeans?”, *Federal Reserve Bank of Minneapolis Quarterly Review*, 28, 2–13.
- ROGERSON, R. and WALLENIUS, J. (2009), “Micro and Macro Elasticities in a Life Cycle Model with Taxes”, *Journal of Economic Theory*, 144, 2277–2292.
- ROGERSON, R. (1988), “Indivisible Labour, Lotteries and Equilibrium”, *Journal of Monetary Economics*, 21, 3–16.