# The Global Life-Cycle Optimizer – Analyzing Fiscal Policy's Potential to Dramatically Distort Labor Supply and Saving*

Johannes Brumm Laurence Kotlikoff

Christopher Krause Joshua Zanger

November 22, 2025

**Abstract**

Fiscal policy in the U.S. and other countries renders intertemporal budgets non-differentiable, non-convex, and discontinuous. Consequently, assessing work and saving responses to policy requires global optimization. This paper develops the Global Life-Cycle Optimizer (GLO), which robustly and precisely locates global optima in highly complex fiscal settings. We use the GLO to study how a stylized U.S. fiscal system distorts workers' labor supply and saving. The system incorporates kinks from federal income tax brackets, Social Security's FICA tax, and a notch from the provision of basic income below a threshold. The GLO reproduces theoretically predicted earnings bunching and flipping over a remarkably wide range of wage rates. Saving distortions and associated excess burdens are substantial. Extensions with discrete labor supply, joint taxation of couples, social security, and labor-income risk demonstrate the versatility of the GLO. The GLO outperforms value function iteration and readily solves cases where value function iteration is infeasible.

*Brumm: Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany, johannes.brumm@gmail.com, corresponding author. Kotlikoff: Boston University, Boston, MA 02215, kotlikoff@gmail.com. Krause: (work done at) KIT, chrkrausecon@gmail.com. Zanger: KIT, j-zanger@web.de. We thank Robert Moffitt, Yongyang Cai, Philip Jung, and seminar participants at Boston University, the European Central Bank (ECB), the Federal Reserve Bank of Atlanta, and at KIT for helpful suggestions. Johannes Brumm and Christopher Krause acknowledge support from the ERC (101042908). Laurence Kotlikoff acknowledges support from the Federal Reserve Bank of Atlanta, Economic Security Planning INC, the Goodman Institute, and Boston University. The authors acknowledge support by the state of Baden-Wuerttemberg through bwHPC.

**Keywords:** Benefits, Distortions, Excess Burden, Global Optimization, Labor Supply, Taxation.
**JEL Classification Codes:** H20, H24, H30, H31, I38, J22

# 1 Introduction

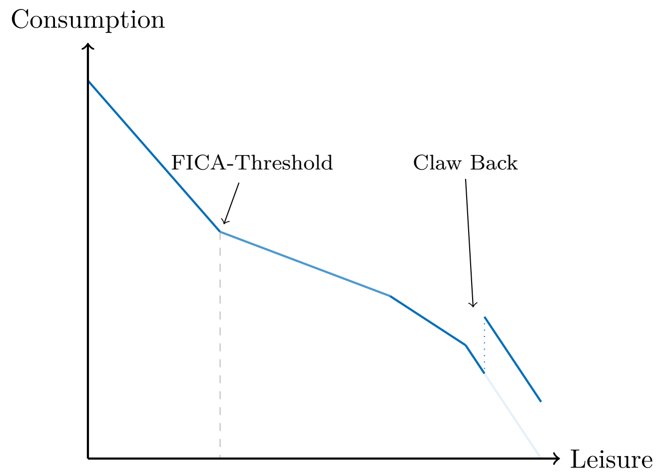
Fiscal systems in many countries are extraordinarily complex due to a plethora of national and state tax and benefit policies. These policies generally come with program-specific eligibility conditions and highly non-linear net payment schedules. Kinks in budget sets from changes in tax brackets are routine. So are notches arising from benefit eligibility and tax thresholds.[1] Furthermore, choice sets are inter-temporally intertwined, not least due to asset-income taxation and the dependence of Social Security benefits on workers' annual earnings histories. The prevalence of highly complex, intertemporal budget constraints is no surprise. Politicians signal their value added by enacting new policies. In so doing, they rarely consider the impact on work or saving incentives, let alone economy-wide efficiency. The US is a case in point. Its fiscal system features over 500 tax and benefit programs comprising federal programs, state programs, and 51 state-specific versions of federal programs.

Understanding the effects of fiscal complexity is challenging. The multitude of fiscal programs renders intertemporal budget sets non-differentiable, non-convex, and discontinuous (NND). In such environments, work and saving responses cannot be assessed by relying on first-order optimality conditions. Instead, they require global optimization to identify the best of all affordable age-consumption and age-leisure paths. But feasible global-solution methods must avoid the *Curse of Dimensionality*, i.e. computation requirements rising exponentially in the dimensionality of the problem – in this case, the large number of continuous choices made over the life cycle.[2] This paper develops the *Global Life-Cycle Optimizer* or GLO to overcome these problems. The GLO searches for globally optimal annual consumption and labor-supply

[1] E.g., U.S. Medicaid, Supplemental Security Income programs, Section-8 housing thresholds, and Medicare Part B IRMAA premium thresholds.

[2] A different form of the Curse of Dimensionality is often encountered in dynamic stochastic economic models, even absent NND constraints, when the number of continuous *states* becomes large (see, e.g., Brumm and Scheidegger, 2017).

Figure 1: Stylized Static Budget Constraint with Kinks and Notch




Note: This static budget constraint includes a kink from the FICA tax, two kinks from changes in income-tax brackets, and a notch from the claw-back threshold for basic income. For illustrative purposes, this figure abstracts from the intertemporal dimension of the problem.

paths building on the pattern search literature (see Torczon, 1997; Audet and Dennis, 2003).

Given the GLO's relative simplicity, its performance is remarkable as we show by applying it to a stylized NND net-tax schedule comprising three elements. The first is a wage tax that includes the seven brackets of the U.S. federal personal income-tax.[3] The second is the 12.4 percent Social Security payroll (FICA) tax levied up to its ceiling.[4] The third element is the provision of $10,000 in basic income for those earning less than $15,000.[5] Figure 1 shows a static NND budget constraint – ignoring saving – including all three elements of our tax scheme.

To demonstrate the GLO's ability to make optimal life-cycle work and saving decisions, we compare the GLO's solution, assuming our NND fiscal system, to that

[3]Our stylized policy taxes, for much of the paper, only tax labor income. This isolates fiscal impacts on labor supply and, given our posited preferences, associated changes to saving. We focus primarily on labor-income taxation for illustrative purposes only. The GLO is fully capable of handling total-income taxation, as we show in section 5.

[4]We ignore, as is standard (see Burtless, 1976), marginal Social Security benefit-tax linkage. This assumption is reasonable given Social Security's immensely complex benefit formulas. The system provides just 12 benefits, but has 2728 basic rules about the 12 benefits in its Handbook.

[5]This notch proxies for the complete loss of Medicaid, Section-8 housing, Supplemental Security Income, and other benefits from earning beyond specified limits. We defer income and asset tests for future analysis.

computed by discrete value function iteration (VFI). We find that VFI and the GLO generate essentially identical solutions, yet the GLO is orders of magnitude faster and more accurate. Moreover, VFI is able to solve the test problem just because its solution requires only one state variable. The GLO, in contrast, can easily be applied – as shown below – to problems that go far beyond the capacity of VFI due to the requisite number of state variables.

Having established the GLO's accuracy, we focus on two goals – demonstrating the GLO's capabilities and versatility and determining labor supply and saving responses to our NND intertemporal budget sets. Our illustrative preferences are separable both across time and between consumption and leisure within a given period. In addition, we posit equal interest and time preference rates to produce flat optimal age-consumption profiles absent asset-income taxation or binding cash-flow constraints. This holds despite the nature of wage taxation. Finally, we often assume a constant wage rate over the life-cycle. This implies constant annual labor supply absent taxation or when taxes are either lump sum or linear in earnings. Consequently, in most of our exercises, the shapes and levels of workers' age labor-supply profiles convey, at a glance, the presence of NND fiscal policy.

Depending on the worker's wage, our modest set of fiscal provisions produces earnings *bunching* and earnings *flipping* over remarkably wide ranges of the earnings distribution. Earnings bunching references earning just below tax kinks – higher marginal tax brackets – and benefit notches – discrete benefit losses arising from exceeding benefit-eligibility thresholds. Earnings flipping entails supplying relatively small amounts of labor in some years and large amounts in others. These policy-induced labor supply reductions dramatically lower annual saving and, consequently, wealth at retirement. Earnings bunching has long been theoretically predicted (see Moffitt, 1983; Burtless, 1976) and empirically documented (see Kotlikoff, 1978; Friedberg, 1998, 2000). Earnings flipping across years is also the theoretically expected response to NND frontiers. It permits workers to partially convexify their lifetime budget sets – to work less and pay lower taxes, on average, over one's workspan.

While the GLO reproduces theoretically-predicted and empirically-documented labor supply behaviors, the wide range of wages over which this occurs is surprising. One would expect major bunching around the $15,000 earnings level to avoid loss of basic income. But it also occurs, for instance, for workers with quite high levels of wages who work less to lower their tax bracket from 32 percent to 24 percent. The reduction in labor supply from earnings bunching can be massive – up to 40 percent in our setting. This difference is relative to labor supply under equal annual lump-sum taxation.

In addition to permanently lowering their labor supply, workers, depending on their wage rates, will flip back and forth between high and low labor supply. Flippers may switch between earning just before the notch in some years and working substantially more in other years. Workers near the FICA kink are also prone to flip. Since the FICA kink is concave, it induces the opposite of bunching, namely a desire to spend part of one's working years working intensively above the Social Security taxable earnings ceiling. The reason is to garner a higher net wage. Another key finding is the extreme sensitivity of life-cycle saving to the fiscal system. Wealth at retirement is, in our worst-case finding, 40 percent lower than it would be with equal-annual lump sum taxes.

We also find moderate to exceptionally large excess burdens measured relative to lump-sum taxation. These values are highly sensitive to workers' wages and fiscal provisions. With all three fiscal elements in place, low-wage workers experience excess burdens, measured as annual consumption equivalents, as high as 28 percent. This reflects the presence of the notch at $15,000 leading, for example, a worker with a full-time wage of $30,000 to earn, annually, only $15,000 in order to keep the $10,000 in basic income. Absent the basic-income-clawback policy, excess burdens are far smaller, ranging between two and just over five percent, with the maximum reached by households earning about $100,000 per year.

To demonstrate GLO's versatility, we also consider discrete labor choice, joint taxation of married couples, internalizing Social Security's benefit formula, and wage-rate risk. Discrete choice references limiting the choice set to working either what we define as full time or what we define as part time in a given year. Doing so can lead to flipping when it would otherwise not arise. It can also eliminate flipping when it would arise were the worker unconstrained. Hence, restricting workers to part- or full-time work is hardly a benign assumption. Taxing married couples jointly has the theoretically predicted (see, e.g., Kaygusuz, 2010; Guner et al., 2012; Bick and Fuchs-Schündeln, 2018) impact of discouraging the relative labor supply of lower-wage spouses. Here, as elsewhere, the GLO delivers both what theory predicts and data confirm. It also provides a quantitative sense of fiscal distortions. Our third extension models Social Security's earnings-based benefit formula, which relates worker's retirement benefits to their past covered earnings history in a highly non-convex and non-differentiable manner. The GLO has no problem identifying, among other things, that optimal labor-supply entails five years of substantially lower labor supply given that only workers' highest 35 years of covered earnings, out of our model's 40-year workspan, enter in their benefit formula. Finally, we illustrate GLO's potential for handling not just deterministic, but also stochastic optimization problems by having GLO maximize lifetime expected utility in the context of uncertain future wage rates.

The paper proceeds as follows. Section 2 reviews the related literature. Section 3 presents our life-cycle optimization problem and details our tax system. Section 4 details the GLO, verifies its accuracy, and highlights its speed. Section 5 first conveys the remarkable responses of labor supply to the kinks and notches of our illustrative fiscal system and then measures the associated excess burden. Section 6 presents the above-listed several extensions. Section 7 summarizes and concludes.

## 2 Related Literature

An early literature documents bunching and other responses to kinks and notches (aka cliffs), uses observed bunching to estimate labor supply elasticities, and questions whether bunching data suffices to identify underlying behavioral parameters.[6] Of most relevance to our paper are studies that recognize the need for global search and implement global search routines. Burtless (1976)'s paper suggested using piece-wise linear budget constraints (PLB) to compare utility along different segments of non-convex budget frontiers. Unfortunately, the PLB approach becomes computationally intractable when expanded to multiple fiscal programs let alone multiple periods. Consequently, Burtless (1976), Friedberg (1998), Friedberg (2000), and others using PLB do so primarily within static (one-period) models featuring a limited number of

[6]This literature includes Kotlikoff (1978), Zabalza et al. (1980), Danzinger and Plotnick (1981), Moffitt (1983), Hausman (1985), Pencavel (1986), Fraker and Moffitt (1988), Rust (1989), Moffitt (1992), Hoynes (1996), Hagstrom (1996), Keane and Moffitt (1998), Blundell and MaCurdy (1999), Eklöf and Sacklén (2000), Meyer and Rosenbaum (2001), Moffitt (2002, 2003), and Blundell and Hoynes (2004). More recent articles, including Saez (2010), Chetty et al. (2011a), Brown (2013), Bastani and Selin (2014), Blomquist et al. (2021), and Bertanha et al. (2023).

fiscal programs. PLB's computational constraints as well as the presence, in the data, of households making theoretically dominated labor supply decisions led MaCurdy et al. (1990) to approximate kinked and notched frontiers with smoothed functions.[7]

Yet, smooth frontiers can't explain earnings bunching or discrete changes in labor supply from year to year. Zabalza et al. (1980), Fraker and Moffitt (1988), and Keane and Moffitt (1998) take a different approach to computing global optima of structural models. They drastically restrict the choice set. Specifically, they assume no saving and restrict labor supply to three choices – no work, part-time work, and full-time work. Blundell and Shephard (2012) examine UK tax reform, including the use of tagging and imperfect observations on hours worked. Theirs is also a static model, but incorporates six different discrete choices of labor supply. Assuming single-period agents facing discrete labor-supply constraints comes at a price. Clearly, most households do save or dis-save. Yes, many households, particularly those with low incomes, are cash-flow constrained. But the degree to which their constraints bind is endogenous to future household choices. Moreover, as we show in section 6, restricting labor supply to discrete options can rule out or rule in suboptimal behavior when such constraints either don't apply or are misspecified.

Like Keane and Moffitt (1998), Rust (1989, 1990) pursues global analysis via discretization. But Rust proposes incorporating intertemporal choice by implementing discrete-state dynamic programming. He also limits consumption and labor supply to discrete values. However, he too readily acknowledges computational limits.[8] These discrete choice studies seek to reconcile actual with modeled behavior. Our goals are different – illustrating the range of potential optimal behavioral responses to NND policy for assumed behavioral parameters and demonstrating the GLO's ability to handle far more complex economic decision problems by avoiding dynamic programming, thereby mitigating the Curse of Dimensionality.

Moore and Pecoraro (2020, 2021, 2023) also restrict labor supply to discrete values in computing global life-cycle optima. But they provide four major improvements over prior discrete-choice analyses. First, they incorporate an extensive range of NND

[7]In so doing, they were able to incorporate wage-rate measurement error, which could empirically reconcile otherwise dominated choices as well as permit testing rationality (Slutsky) conditions.

[8]For example, fully considering Social Security requires treating all past covered earnings as state variables. Rust sidesteps this problem by assuming Social Security benefits are determined by average past-covered earnings.

fiscal policies. Second, their analysis is general equilibrium, entailing computation of the transition path of life-cycle economies featuring NND policies. Third, they incorporate a wide range of realistic elements, including housing choice, the decision to rent or own, fixed costs of working, home production, child care costs, bequests, uncertain longevity, and more. And fourth, building on Carroll (2006) and Iskhakov et al. (2017), they introduce a hybrid endogenous grid method approach to dynamic programming that avoids interpolation, permits global optimization, yet suffers from the curse of dimensionality in the number of state variables. Moore and Pecoraro's papers teach valuable lessons. The most important for our analysis is their finding that incorporating explicit NND fiscal policy materially matters both to microeconomic behavior and macroeconomic outcomes.

Before developing the GLO as a robust global optimizer for deterministic NND life-cycle models with continuous choice, we tried a variety of traditional global optimization methods, including genetic and random search algorithms – without success. Building on the pattern search literature (see Torczon, 1997; Audet and Dennis, 2003) and including features tailored to fiscally realistic life-cycle problems proved successful. Of course, other methods, including those discussed and developed in Arnoud et al. (2019) and Guvenen (2011), may be able to match GLO's optimization performance if properly adapted to handle such fiscal conditions. Arnoud et al. (2019) provide a comparison of specific global optimizers, including several versions of TikTak developed in Guvenen (2011), which they apply to method-of-simulated-moments estimation as well as to analytical test functions. We also test the GLO against these functions, in appendix C, highlighting its speed, reliability, and scalability.

One might also ask whether neural nets could help find global solutions in our setting. Azinovic et al. (2022), Maliar et al. (2021), and Duarte et al. (2021) demonstrate the impressive power of machine learning to handle complex life-cycle problems, including various forms of uncertainty. However, as discussed in Duarte et al. (2021), machine learning is ill suited to deal with NND frontiers due to its reliance on differentiability.

# 3 The Life-Cycle Problem

Our life-cycle model is simple apart from assumed kinks and notches in the net tax schedule. Depending on the case under consideration, income will reference either wage income or total (wage plus asset) income.

## 3.1 Lifetime Utility

Households live for $T$ periods with lifetime utility given by

$$\sum_{t=1}^{T} \left( \frac{1}{1+\rho_t} \right)^{t-1} U(c_t, l_t), \tag{1}$$

where $\rho_t$ is the time preference rate, $c_t$ is consumption, and $l_t$ is labor supply, all at time $t$. We assume that per-period (annual) utility obeys the commonly-applied King et al. (1988) additively separable functional form,

$$U(c, l) = \log c - \chi \frac{l^{1+1/\gamma}}{1+1/\gamma}, \tag{2}$$

where $\gamma$ is the Frisch elasticity and $\chi$ is a scaling parameter, which is set to 1 for most of the analysis. Households work for the first $R$ periods only, thus $l_t = 0$ for $t > R$. The per-period budget constraint in $t$ is:

$$c_t + a_{t+1} = w_t l_t + (1 + r_t) a_t - \mathcal{T}(y_t), \tag{3}$$

where $a_t$ are the beginning-of-period assets, $w_t$ is the wage rate, $r_t$ is the interest rate on assets, $\mathcal{T}$ is the net-tax function, and $y_t = w_t l_t$ is labor income.[9] Note that if the after-tax wage rate is constant and the interest rate equals the time preference rate, $r_t = \rho_t$, labor supply is constant. We focus on examples satisfying this assumption since it implies that any deviation from a constant age labor-supply profile reflects the kinks or notches of our net tax schedule.

Throughout the paper, we set the period length to be one year, and for the most part, we make the following parameter choices. We set the annual interest rate on

[9] We consider taxing total income, i.e. $y_t = w_t l_t + r_t a_t$, in section 5.

Table 1: Tax Brackets and Marginal Tax Rates

| Taxable Income | Marginal Tax Rate |
|---|---|
| $\$0 - \$10,275$ | 10% |
| $\$10,275 - \$41,775$ | 12% |
| $\$41,775 - \$89,075$ | 22% |
| $\$89,075 - \$170,050$ | 24% |
| $\$170,050 - \$215,950$ | 32% |
| $\$215,950 - \$539,000$ | 35% |
| $539,000+ | 37% |

Note: Tax brackets for single filers in 2022.

assets and the time preference rate to 2 percent. For the Frisch elasticity, we choose a value of 1, which is among the higher range of values found in the empirical literature (see Chetty et al., 2011b). We assume that household have a 40-year workspan and a 60-year lifespan. Finally, agents are not able to borrow, that is $a_{t+1} \geq 0$ for all $t$.

## 3.2 Stylized Tax System

We consider net-tax codes with $n$ tax brackets, each characterized by cutoff value $b_i$, intercept $x_i$, and proportional tax rate $\tau_i$:

$$\mathcal{T}(y_t) = \begin{cases} x_1 + t_1 y_t & \text{if } y_t \in [0, b_1] \\ x_i + t_i(y_t - b_{i-1}) & \text{if } y_t \in ]b_{i-1}, b_i], \quad \text{for } i = 2, \dots, n, \\ x_n + t_{n+1}(y_t - b_{n-1}) & \text{if } y_t \in ]b_n, \infty[. \end{cases}$$

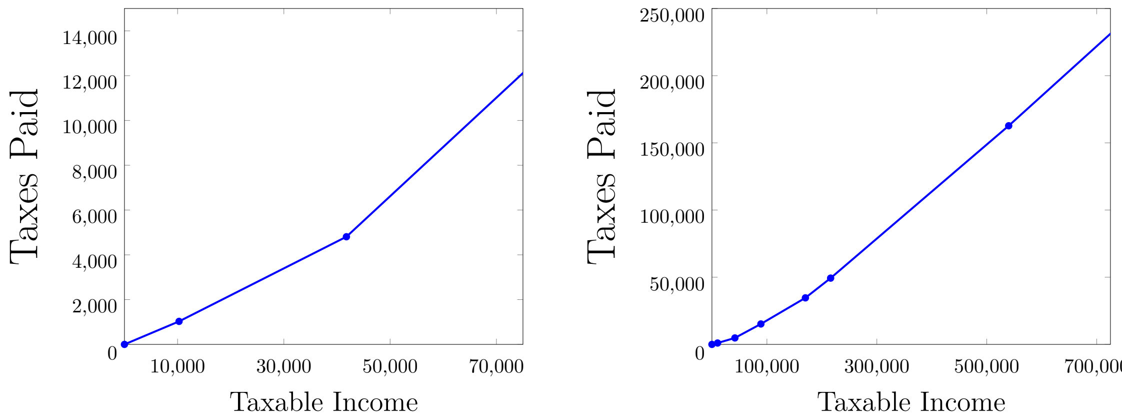
Our baseline net-tax schedule comprises three components. First, the 2022 US single's tax brackets reported in table 1. Second, Social Security's 12.4 percent Social Security FICA tax on labor earnings through \$147,000.[10] Third, we posit a basic income of \$10,000 paid to those with incomes below \$15,000 – a proxy for welfare-benefit thresholds arising under the Supplemental Security Income, Medicaid, Section-8 Housing,

[10]There is no ceiling on the 2.73 percent Medicare FICA tax. In what follows, the FICA tax denotes only the Social Security portion of the overall FICA tax. We also treat the division between "employer" and "employee" portions of the FICA tax as economically irrelevant; i.e., we assume workers bear the full tax regardless of how the remittance of the full tax by employers to the government is labeled.

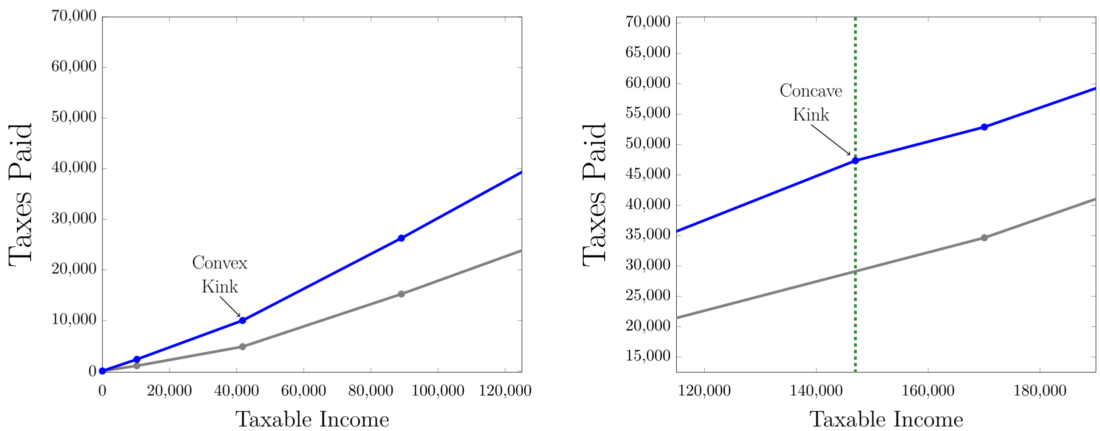
and other programs. In addition to considering this baseline schedule, we also present results assuming just income taxation, $\mathcal{T}_I$, and just income and FICA taxation, $\mathcal{T}_{IF}$. When comparing the three schedules we denote the baseline tax schedule by $\mathcal{T}_{IFB}$ and refer to the three schedules as "INC", "INC+FICA", and "INC+FICA+BASIC". Figure 2(b) shows that INC has convex kinks only and that the FICA tax adds a non-convex kink.[11] Finally, figure 2(c) displays the notch at the claw-back threshold for basic income present in our baseline tax code INC+FICA+BASIC.

[11]Note that the convex (concave) kinks in the tax schemes displayed in figures 2(a) and 2(b) correspond to concave (convex) kinks in budget sets as depicted in Figure 1.

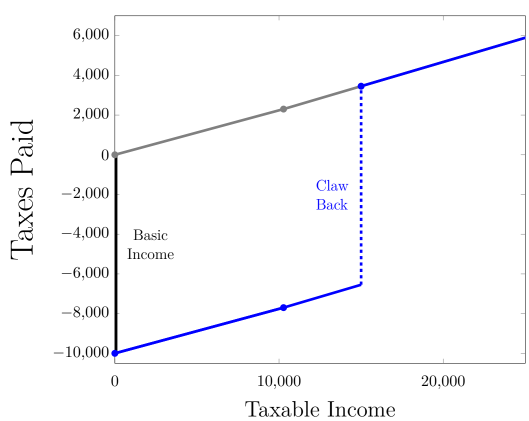
Figure 2: US Tax Code: Kinks and Notch




(a) Income Tax – Convex Kinks Only




(b) Income Plus FICA Tax – Convex Kinks and one Concave Kink




(c) Income Plus FICA Tax Code With Basic Income – Notch at Claw-Back Threshold

# 4 The Global Life-Cycle Optimizer

This section first presents the GLO algorithm and then showcases its ability to solve highly NND life-cycle problems efficiently and accurately.

## 4.1 The Algorithm

The GLO algorithm is a variant of pattern search – a well established global optimization method.[12] The basic algorithm seeks the global maximum of a function $f : \mathbb{R}^N \to \mathbb{R}$. The GLO starts from an arbitrary point, $x^0 \in \mathbb{R}^N$. To generate the next point in the sequence, $x_{k+1}$, the algorithm carries out a so-called polling step.[13] In each polling step a set of $J$ points, the poll set, is constructed by moving from the current point in $J$ different directions, i.e. adding $J$ different vectors to $x_k$. Each of these $J$ vectors is given by multiplying a spanning direction $d \in D = \{d^1, \ldots, d^J\}$ with a scalar $m$, called the mesh size.[14] The poll set is thus given by:

$$\mathcal{P} = \{x_k + m \cdot d : d \in D\} \tag{4}$$

If one of the points in the poll set, $p \in \mathcal{P}$, improves the objective function relative to the current point, $f(p) > f(x_k)$, the poll is called successful and the point that achieves this improvement becomes the current point in the next iteration, $x_{k+1} = p$.[15] If no improvement is found, $f(p) \leq f(x_k)$ for all $p \in \mathcal{P}$, then the current point is retained for the next iteration, $x_{k+1} = x_k$. Furthermore, if the poll was successful, the mesh size is increased such that the algorithm "zooms out" and considers a larger space to find further improvements. In the case of an unsuccessful poll, the mesh size is reduced, i.e., it "zooms in." The procedure continues until the mesh size falls below a given threshold.

[12] See Torczon (1997) or Audet and Dennis (2003) for a general exposition.

[13] The general pattern search algorithm can also include a so-called search step in addition to the polling step, see Audet and Dennis (2003).

[14] A common pattern for the spanning directions is to vary only one dimension at a time, yet in both directions, thus setting $J = 2N$ with $D = \{(1, 0, \ldots, 0), ..., (0, 0, \ldots, 1), (-1, 0, \ldots, 0), ..., (0, 0, \ldots, -1)\}$.

[15] If one chooses a complete polling, the point associated with the best objective function value is chosen. Alternatively, one could stop the polling after any improvement is found and go to the next iteration.

Our life-cycle problem consists of finding the agent's paths of consumption and labor supply that maximize lifetime utility, or, as illustrated in an extension, lifetime expected utility, subject to per-period cash-flow constraints. To apply pattern search to this specific optimization problem, we make three substantial adjustments to the basic algorithm. First, when constructing elements of the poll set we simultaneously consider changes in more than one dimension. We draw two of these dimensions randomly, permitting trade offs between any two values of consumption and/or labor supply. In addition, we always adjust terminal consumption to ensure lifetime budget balance, that is, all remaining assets are consumed in the last period. The second modification makes the change in one of the chosen dimensions stochastic, replacing the mesh size with a random number drawn from a uniform distribution on an interval around the mesh size. Our third adjustment ensures that all points in the poll set satisfy the cash-flow constraints. For each point, we successively consider each period's budget, starting at $t = 1$; in case the cash-flow constraint is violated, we adjust consumption so that the constraint is precisely met. With this combination of features, GLO's algorithm can, as now shown, tackle highly NND life-cycle problems.

Appendix B provides a formal description of the GLO algorithm. Our application of the GLO to our specific life-cycle NND problem proceeds as follows. Recall that agents work from 25 to 65 years and consume from 25 to 85 years. Consequently, their optimal solution comprises a 100-element vector with 60 annual levels of consumption and 40 annual levels of labor supply. The GLO starts by setting a mesh size and choosing a (random) starting guess. The 100th element references age-85 consumption, which is always set to satisfy the household's intertemporal budget given values of the other elements. Call this vector $x$ and denote the resulting lifetime utility $f(x)$. The GLO next chooses two of the first 99 elements at random. The first of the two values is both increased and decreased by the mesh size. The second of the two values is both increased and decreased by the current mesh size multiplied by a number randomly chosen from a uniform distribution between 0 and 1.25. The four up-up, up-down, down-up, down-up perturbations together with the other 96 vector elements (including the lifetime-budget balancing final consumption level) provide four candidate solution vectors. We evaluate our objective function for each and repeat the process 500 times. We then find the maximum of the 2000 (4 times 500) evaluations. Call the maximizing vector $p$. If $f(p) > f(x)$, set $x = p$ and double the mesh size. If $f(p) < f(x)$, leave $x$ unchanged and reduce the mesh size in half. Next, we repeat the above process starting with the potentially revised $x$ until the mesh size falls below a specified value. Finally, we restart the entire algorithm several times (depending on the complexity of the problem) with either different starting guesses and/or different seeds for the random number generator, and take the maximum of the stored optima as our solution.

Why does the GLO find the global optimum? Intuitively, if the candidate solution vector is far from the optimum, the algorithm will move to a vector that is close to a local or global optimum. Near a local optimum, large and small deviations (by the mesh size) may succeed. Small deviations that succeed result in an increase in the mesh size and thus the subsequent step may move the solution away from the local optimum. Near a global optimum, small deviations may succeed, bringing the solution even closer to the global optimum. Large deviations, in turn, will be unsuccessful resulting in an ever smaller mesh size eventually triggering the convergence criterion, which ends the algorithm. The combination of initializing the GLO with a random guess and incorporating multi-starts delivers the global optimum with, as now shown, high accuracy.

## 4.2 Testing the Algorithm

For the class of problems we aim to solve, with NND intertemporal budget sets, analytic solutions are not available to test GLO's accuracy and its convergence behavior. To convince ourselves of the reliability of the GLO algorithm, we, therefore, proceed in two steps. First, we apply the GLO to solving standard analytical test functions for global optimization. We pick functions that Arnoud et al. (2019) use to solve up to 10-dimensional problems and find that, as appendix C reports, the GLO can quickly, precisely, and robustly find the analytically known global optimum of these test functions in up to 100-dimensional problems. Second, we compare the GLO's results on our NND life-cycle problems with the results of discrete value function iteration (VFI), which is in principle able to solve these problems, albeit with enormous costs and subject to a very strong Curse of Dimensionality. We find that for our test-problem with only one state variable, namely asset holdings, VFI and the GLO basically identify the same solution, yet the GLO is orders of magnitudes faster and, indeed, more accurate. Note, however, that the GLO can easily be applied to

Table 2: GLO versus Value Function Iteration

| | Wage Rate | | | |
|---|---|---|---|---|
| | **\$40,000** | **\$80,000** | **\$130,000** | **\$200,000** |
| **CEV (GLO vs. VFI)** | | | | |
| Best | +0.0159% | +0.0175% | +0.0174% | +0.0175% |
| Median | −0.0850% | +0.0175% | +0.0172% | +0.0175% |
| **Runtime (s)** | | | | |
| GLO | 123 | 55 | 55 | 54 |
| VFI | 11132 | 11495 | 11388 | 11769 |

Notes: Differences between the GLO solution and VFI solution, measured in CEV, show that the (best) GLO solution is slightly better than the VFI solution based on the value of the objective function. This confirms GLO's ability to identify global maxima of NND problems.

problems with large numbers of state variables, as in the Social Security example of section 6.3, whereas VFI cannot.

For our comparison between the GLO and VFI, we solve the life-cycle problem of section 3 with different wage rates. Appendix C details our VFI implementation. But to summarize, we parallelize the VFI on 80 cores and also use 80 cores for the GLO by letting it run 80 times with differing starting guesses. To compare the solutions found by VFI and the GLO, we compute the percentage deviation in consumption equivalent variation (CEV), as defined in appendix A, between the VFI solution and the different GLO solutions. Table 2 reports these deviations for both the best (among 80) solutions of the GLO and the median solution of the GLO. It also reports run times.

We find that the best GLO solution is superior to the VFI solution for all wage rates – which are held constant over the life-cycle for simplicity, although a non-constant profile poses no problem for the GLO as shown below. The median GLO solution is basically as good as the best VFI solution for all wage rates except \$40,000 – a low enough wage to make the basic-income notch relevant. But even in this case, the median solution deviates from the best solution by just about one tenth of one percent.[16]

[16]This observation justifies adjusting GLO's algorithm to let it run to completion based on numerous initial guesses and then picking the best solution – a multi-start routine as is common in global optimization.

As for compute time, the GLO is extremely fast, solving the hardest problem in roughly two minutes and solving all other problems in roughly one minute. VFI in turn, takes about three hours on all problems, despite being parallelized and optimized in various ways as explained in Appendix C. The main takeaway here, however, is not that the GLO is faster than VFI's brute force method in solving NND problems. Instead, it's the GLO's ability to derive essentially the same solution as VFI, certifying its ability to identify global maxima of NND problems. Many problems, including the Social Security example of section 6.3, are not substantially harder for the GLO to solve than the test problem at hand, but could never be addressed with VFI due to the Curse of Dimensionality.

# 5 Labor Supply Response to Kinks and Notches

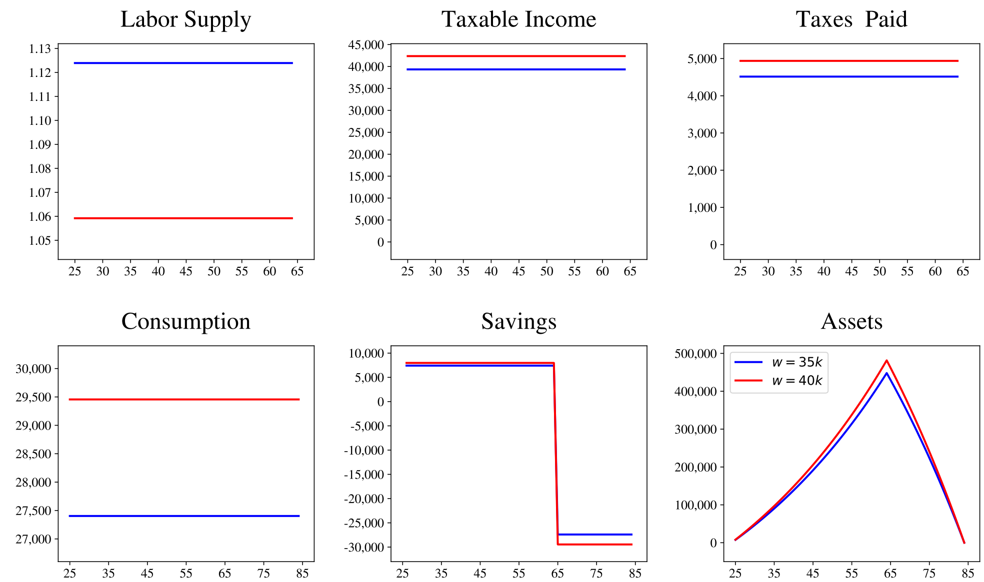
In this section, we first analyze labor supply bunching as a response to convex kinks in the income-tax code leaving out FICA as well as the provision of basic income. Second, we add the FICA tax, which introduces a concave kink and thereby generates flipping behavior. Third, we examine the combination of flipping and bunching arising from including basic income and its associated notch. Finally, we discuss the distortions and excess burden that are associated with all these elements fo the tax code. Most of this section assumes worker's wages are fixed in real terms throughout their working years.

## 5.1 Bunching at Convex Kinks From Income-Tax Brackets

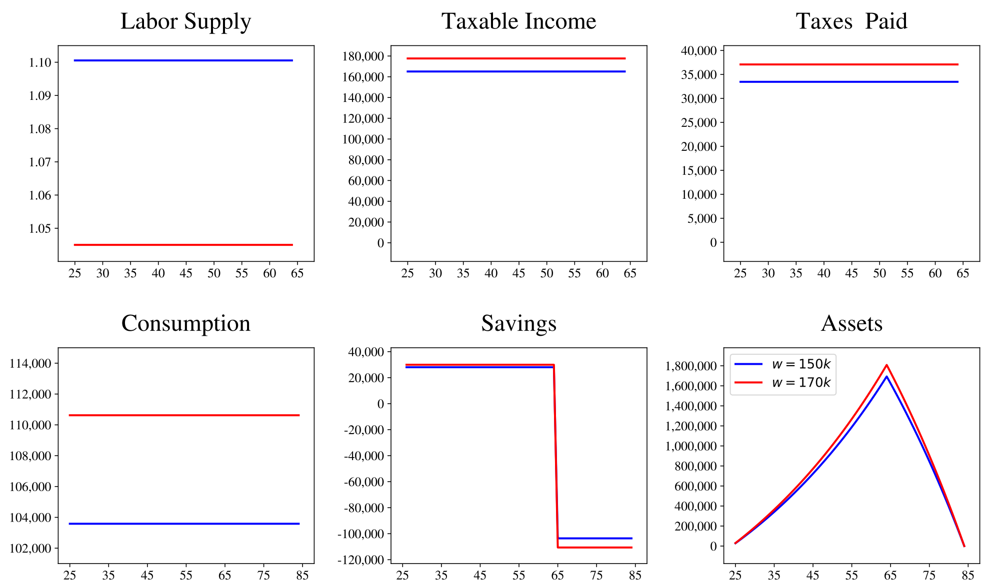
We start by focusing on the two income-tax kinks that are most convex – the jump in the marginal tax rate from 12 to 22 percent occurring at $41,775 and the jump from 24 to 32 percent occurring at $170,050. Consider two workers earning wage rates of $35,000 and $40,000 for one unit of labor supply (i.e., one year of full-time work). As figure 3(a) shows, the lower-wage worker works substantially more than the higher-wage worker.[17] As detailed below, this reflects the high-wage worker's response to her much higher marginal tax rate. As a consequence, the higher-wage

[17]If we consider one unit of labor as the average number of hours worked in the U.S. in 2022, then a difference of five percentage points represents about 90 hours of work.

Figure 3: Life-Cycle Profiles of Households Facing Tax Code with Convex Kinks



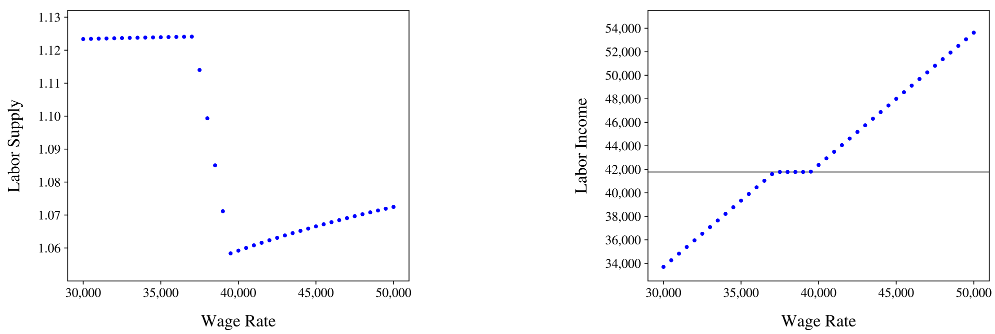

(a) Low-Wage-Rate Households



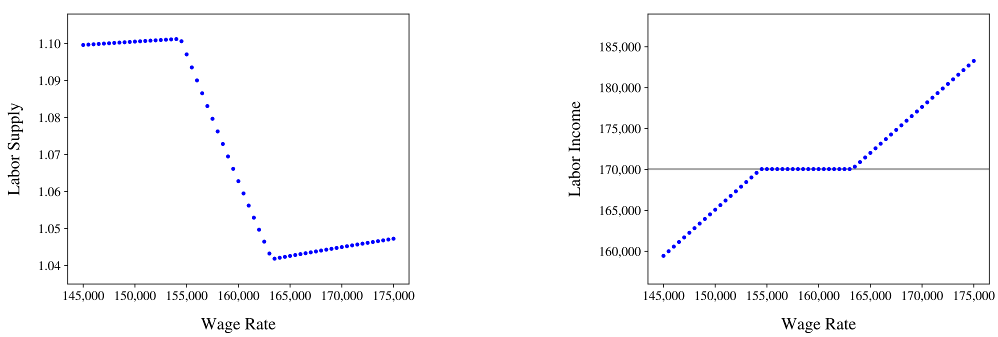

(b) High-Wage-Rate Households

Note: Life-cycle profiles for households facing labor taxation under the income tax code with convex kinks only (INC tax). Consumption profiles are flat as interest rate and time preference rate are equal and asset income is not taxed. Both assumptions are for illustration and are relaxed below.

Figure 4: Bunching at Convex Kinks




(a) Convex Kink at $41,775



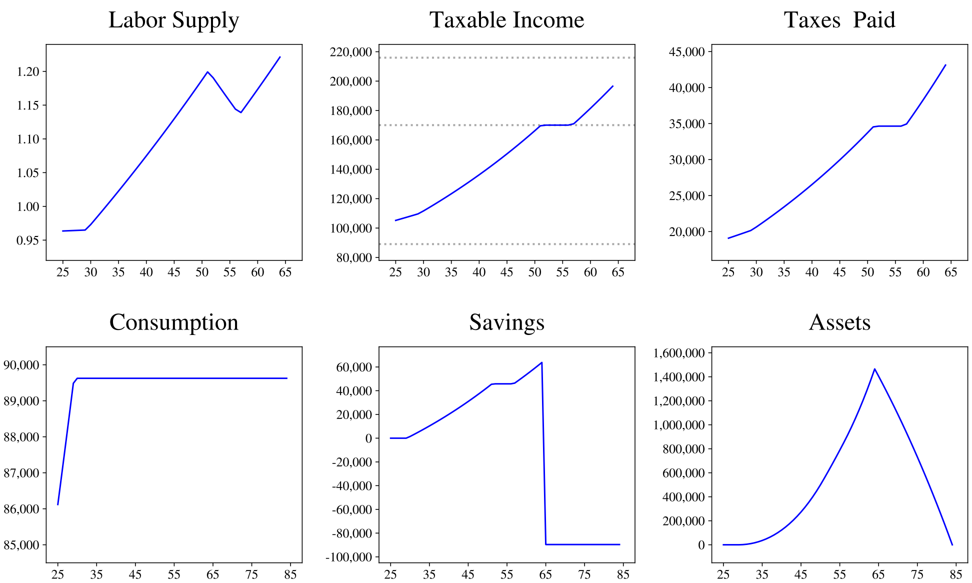

(b) Convex Kink at $170,050

Note: Labor supply and labor income as a function of the wage rate. Bunching occurs on the gray lines, which represent boundaries between two tax brackets.

worker's consumption is only 7.5 percent higher while earning 14.3 percent more per unit of labor. Figure 3(b) provides a similar picture for the case of high-wage-rate households in the vicinity of the convex kink at $170,050.

Next, we consider the labor supply behavior for ranges of wage rates around the two highly convex kinks of the income tax code. Figure 4(a) shows how labor supply and labor income depend on the wage rate. For a substantial range of wage rates – from $37,500 to $39,500 – households chose to earn exactly $41,775, the amount corresponding to the major convex kink at the lower end of the tax schedule. Within this range of wage rates, labor supply is strongly decreasing in the pre-tax wage. Starting at a wage of about $40,000, households are willing to pay the higher marginal tax rate, and labor supply and income both increase. However, labor supply remains

Figure 5: Bunching over the Life-Cycle when the Wage Rate Grows




Note: Life-Cycle profiles of high-wage-rate household with one percent yearly wage growth.

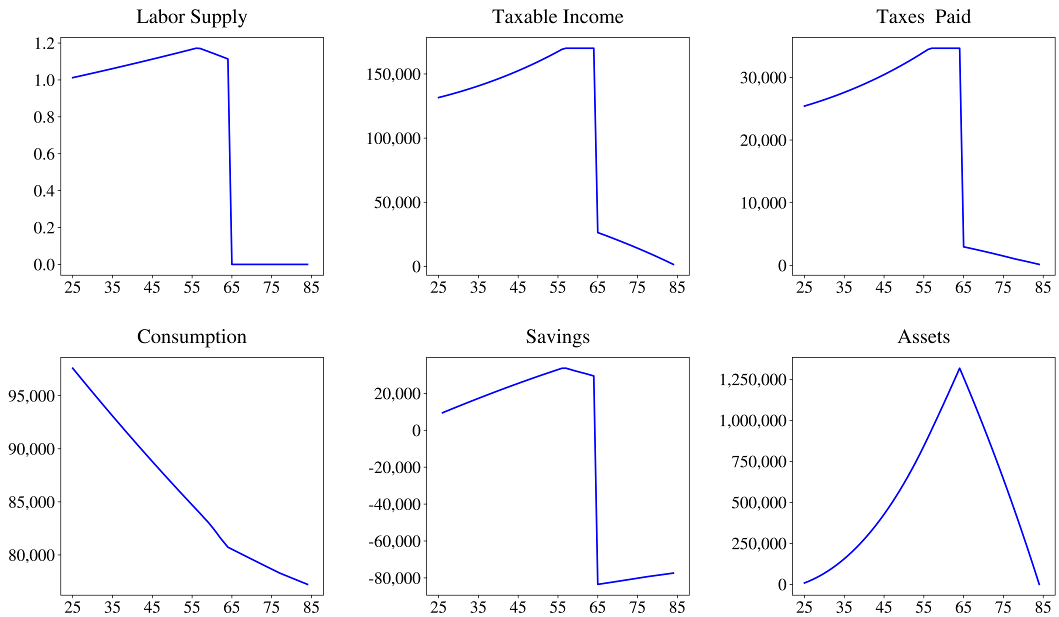
substantially below the labor supply of agents with lower wage rates. As shown in figure 4(b), the behavior is qualitatively similar at the other highly convex kink – $170,050. The range of bunching wage rates, however, is much larger, ranging between $154,000 and $163,000.

We now turn to bunching across time. Consider, in figure 5, bunching by a worker whose wage rate rises by one percent annually starting at $109,184.[18] This worker's labor supply remains low for several years. Then it begins to rise, but not by enough to push the worker above the 24 percent bracket. This continues until the worker reaches the $170,050 bracket threshold. At this point, the worker reduces their labor supply each year for several years to avoid moving into the 32 percent bracket, which they eventually find optimal to do – when their wage becomes sufficiently high.

Bunching across time can also occur with a fixed wage rate if total income is taxed. So far, we've focused on labor-income taxation to isolate its impact on labor supply. We now apply the federal tax schedule to total income, including asset income. Figure 6 displays life-cycle profiles of a household with a wage rate of $130,000.

[18]Note that the present value of this wage path is equal to the present value of earning a constant annual wage of $130,000.

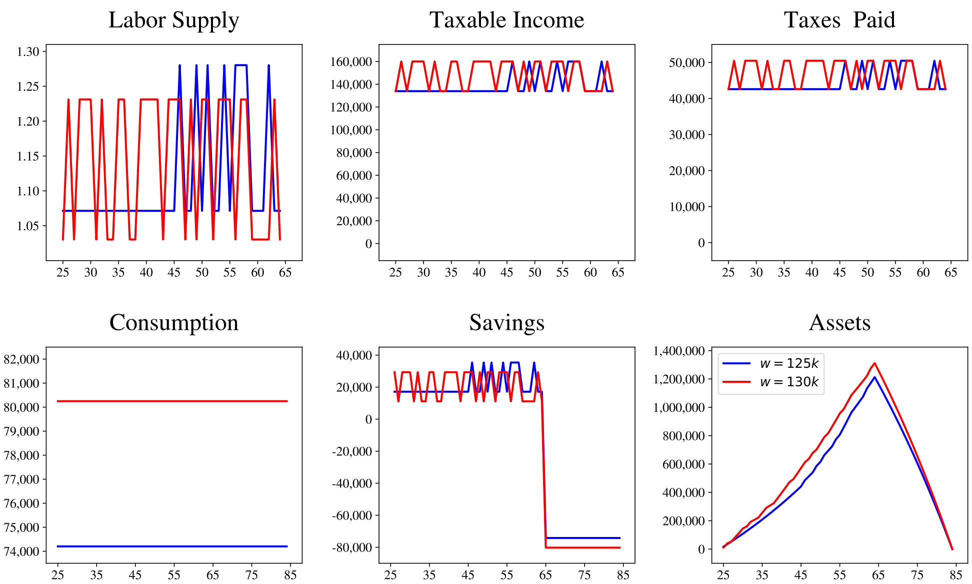
Figure 6: Life-Cycle Profiles of Household Facing Total Income Taxation




Note: Life-Cycle profiles of a household with a wage rate of $130,000 when total income is taxed.

The consumption profile is now generally downward sloping as capital-income taxation discourages saving. Stated differently, as theoretically expected, taxing saving (actually, the return to saving) encourages current over future consumption. Similarly, labor supply tends to be upward sloping since capital-income taxation favors working more when old (taking leisure earlier). However, the convex kink at $170,050 interferes with this pattern. As the household accumulates assets and increases its labor supply, taxable income grows over the life-cycle until it reaches $170,050. From then on, the household reduces its labor supply and adjusts its saving to stay just at that level of total income – the household bunches over time at a total income corresponding to the upper bound of a tax bracket. All in all, the above analysis shows that the US tax code has a strong tendency to induce bunching and that the GLO is able to robustly solve the resulting optimization problem with NND budget sets.

Figure 7: Life-Cycle Profiles of Households Facing Income Plus FICA Tax



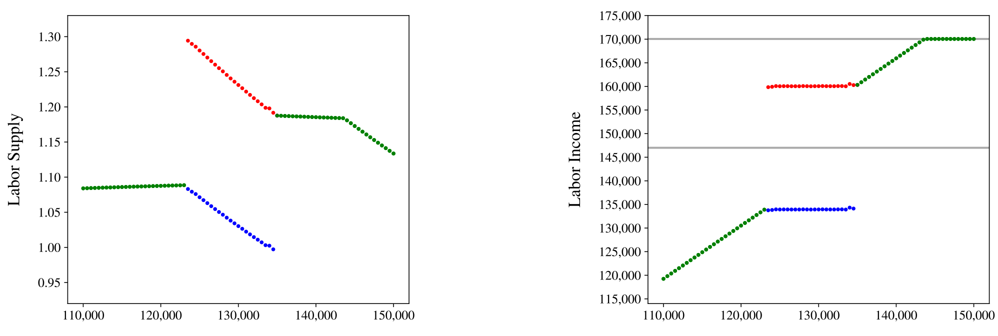

Note: Life-Cycle profiles of household with wage rates close to the FICA threshold of $147,000 exhibiting flipping around that threshold.

## 5.2 Flipping due to FICA's Concave Kink

We next include the 12.4 percent FICA payroll tax with its ceiling (concave kink) at $147,000. Figure 7 displays life-cycle profiles for households with annual wage rates of $125,000 and $130,000. Despite having flat consumption profiles, their labor supply profiles exhibit frequent flipping. In particular, agents jump back and forth between two tax brackets – below and above the FICA threshold of $147,000. Although the lower bracket is characterized by a higher marginal tax rate, continually working sufficient hours to exceed the FICA threshold in all periods is sub-optimal. At the same time, it's also sub-optimal to always stay below the threshold. Hence, the optimum entails working less in some years and more in others.

Figure 8 shows how labor supply and, thus, labor income are affected by the non-convex kink. Workers with wages between $123,000 and $134,500 flip. To be precise, they flip between low labor supply (the blue dots in the plot) which generates income substantially below the FICA threshold (gray line) and high labor supply (the red dots) which generates income much higher than the FICA threshold. Generating

Figure 8: Flipping at FICA's Concave Kink




Note: Labor supply and labor income as a function of the wage rate. Flipping occurs in the neighborhood of the lower gray line that represents the FICA threshold at $147,000. Households earning between $123,000 and $134,500 flip between the blue and the red points. Households with lower or higher wage rates do not flip, corresponding to the green points.

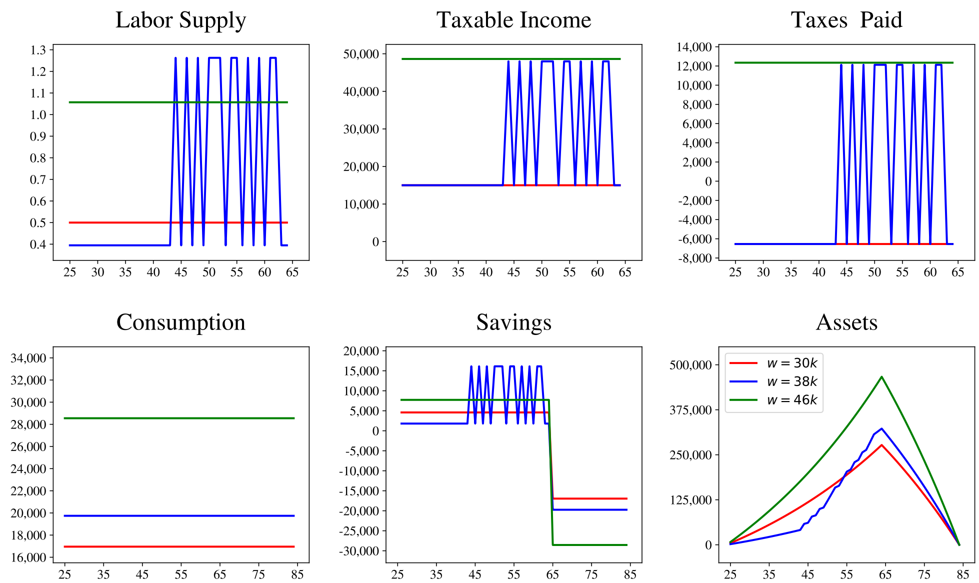
labor income close to the FICA threshold is never optimal – the exact opposite of bunching.

## 5.3 Basic-Income Claw-Back Notch: Bunching and Flipping

We now turn to flipping and bunching due to the notch at $15,000 in the tax and transfer scheme. Recall, the notch is generated by the assumption of a $10,000 basic income that (single) households receive provided they earn less than $15,000. Figure 9 displays the age-labor supply profile for workers with three different wage rates – $30,000, $38,000, and $46,000. Figure 10 displays the labor supply and resulting labor income for a wide range of wage rates.

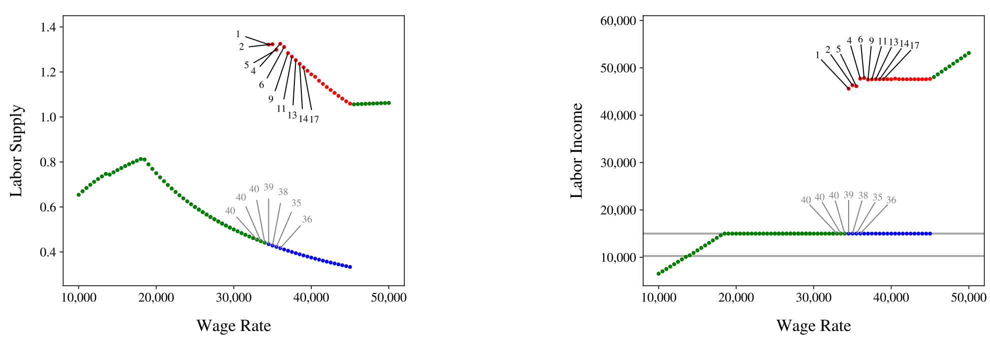
Workers with a wage rate below roughly $35,000 reduce their labor supply each year to stay below the $15,000 threshold. Starting from wage rates around $45,500, individuals choose to never collect the basic income. Between these values, however, it's optimal to collect basic income in some years and to earn substantially above the threshold in other years. In fact, it is never optimal to supply labor that generates anything strictly between the $15,000 threshold for the basic benefit and the upper bound for the 12 percent tax bracket at $41,775. Households who flip accumulate substantial savings in the periods they work. While one might argue that this is ruled out by asset tests in most real-world transfer schemes, we believe that such

Figure 9: Life-Cycle Profiles of Households Facing Basic-Income Claw-Back Notch




Note: Life-cycle profiles for three (low) wage rates, one resulting in below-threshold labor supply throughout working life, one in above-threshold labor supply, while the middle one implies flipping.

Figure 10: Bunching and Flipping at the Basic-Income Claw-Back Notch




Note: Labor supply and labor income as a function of the wage rate. Between \$19,000 and \$45,000 the notch induces bunching. In addition, between \$35,000 and \$45,000, there is flipping between blue and red points. Numbers indicate in how many periods the respective choice is made.

behavior is still relevant. There are often exceptions for assets, like housing and cars. In addition, households can hide assets by 'gifting' them to friends or relatives.

## 5.4 Distortions of Labor Supply and Excess Burdens

Figure 11(a) displays our fiscal system's impact on labor supply compared with that arising under lump-sum taxation. In considering the results, bear in mind that, given our preferences, a simple linear tax would make no difference to labor supply. When households face just the income-tax schedule with its convex kinks, labor supply is reduced by five to sixteen percent and this distortion tends to increase in the wage rate. Intuitively, non-linearity accentuates substitution effects. When the FICA tax is added to the schedule, labor supply is distorted substantially more, especially for those with low wage rates.[19] Interestingly, the labor distortion is now no longer monotonically increasing in the wage rate, as marginal tax rates are no longer monotone. Finally, when we add basic income and its clawback at $15,000, labor supply is massively distorted for low-wage households.

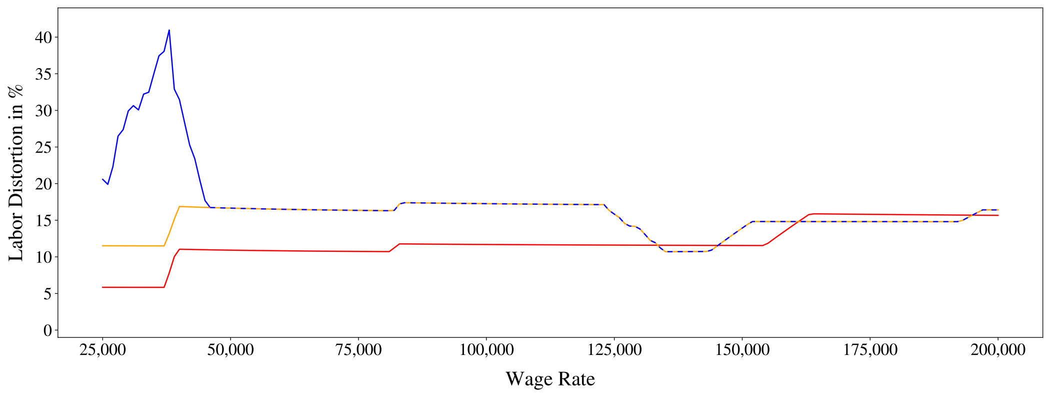
Of course, the labor supply distortions just described have a substantial impact on savings. Figure 11(b) reports the percentage deviation of wealth at retirement relative to the corresponding lump-sum tax case. In our baseline tax scheme (blue line), wealth accumulation is reduced substantially for all wage rates, with reductions ranging between 15% and 40%. Accumulated wealth inherits the non-monotonicity from the labor supply results. The largest reduction in savings is experienced by the low-wage rate workers, while the smallest reduction is experienced by households close to the FICA threshold.

Analysis of the excess burden from taxation – the cost of distorting household consumption and leisure (labor supply and saving) decisions – was made a mainstay of public finance by Harberger (1964).[20] We build on these traditional studies by measuring excess burdens when fiscal systems are NND. To measure the distortions produced by our tax system, we follow the standard procedure – compare, as a consumption equivalent variation (CEV), distorted lifetime utility with lifetime utility under lump-sum taxation. CEV measures the percentage increase in annual con-

[19]The FICA tax is proportional. For our preferences, proportional taxes have no impact on uncompensated labor supply. But the distortion we measure involves compensated labor supply, i.e., just substitution effects, since we are "compensating" workers for the elimination of distortionary taxation with lump sum taxation. Hence, raising total marginal taxation, even proportionately, exacerbates the distortion.

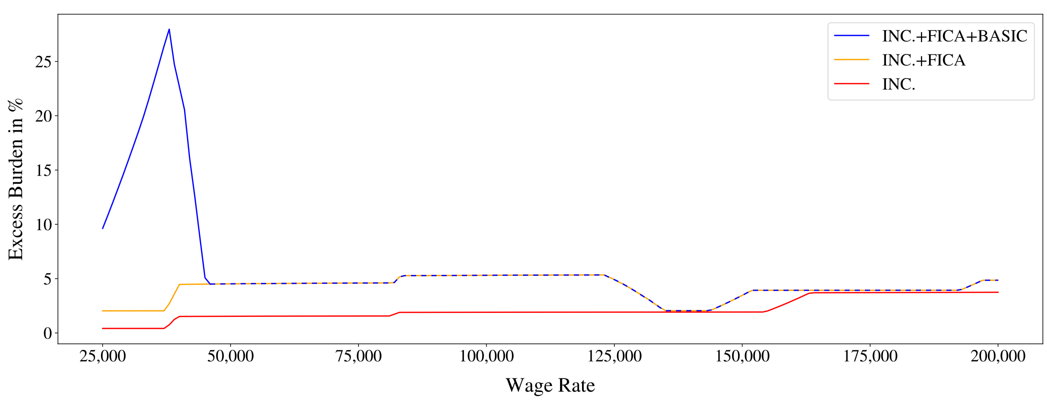
[20]A multitude of studies applied Harberger's approach to all manner of tax-induced distortions. Auerbach (1985) and Auerbach and Hines (2002) provide extensive reviews of this literature.

Figure 11: Distortions and Excess Burdens




(a) Labor Supply Distortion


(b) Impact on Wealth Accumulation




(c) Excess Burden of Different Tax Schemes

Note: Comparison of the different stylized tax schemes with lump-sum taxation. Panel (a) shows labor supply distortion and panel (b) distortion of saving for retirement. Panel (c) reports the excess burden of the respective tax schemes measures in CEV.

sumption that will raise a distorted worker's lifetime utility to that enjoyed under lump-sum taxation. Appendix A provides the precise formula. To control for cash-flow constraints, we collect, from each worker at a given age, the same lump-sum net taxes as they pay when facing one of three distorted tax systems. The first includes just the federal income-tax kinks, the second includes those kinks, plus the concave kink from the FICA tax, and the third includes all kinks plus basic income and the notch from its clawback.

Figure 11(c) shows that the full set of tax provisions produces huge distortions for low-wage workers. The distortion of labor supply and saving decisions peaks around 40 percent and the associated excess burden at over 25 percent. The massive fiscal burden arising solely from distorting behavior drops dramatically for workers earning $50,000 or more, for whom the basic income provision becomes irrelevant. As the red line shows, federal income taxes produce, on their own, small excess burdens apart from those earning above $160,000 pre-tax. When we add in the FICA tax, it becomes moderate, at about five percent, for medium wage households.

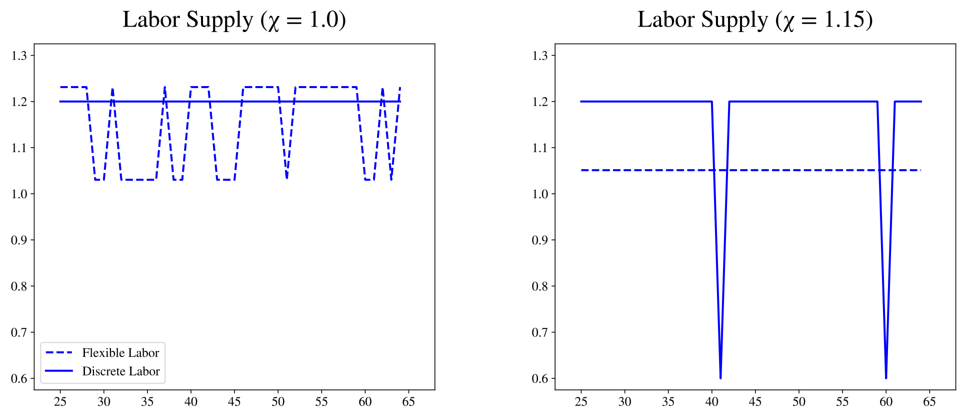
# 6 Discrete Choice, Couples, Pensions, and Risk

Can the GLO handle discrete choice, joint taxation of spouses, and social security? Yes, as this section shows. This is expected. Given GLO's solution method, there is nothing special about these cases or, indeed, incorporating additional elements, such as a distaste for participation in welfare programs or other fiscal provisions emphasized by Moffitt (1983). Including life-cycle risk, in contrast, poses a much greater challenge for the GLO, which we address in our final extension.

## 6.1 Discrete Labor Supply

Next, we compare optimal choice for the hitherto considered case of fully flexible labor-supply with the case of discrete choice. Discrete choice references limiting the choice set to working either what we define as full time or what we define as part time in a given year. As our example shows, restricting behavior to specific discrete choices when such restrictions don't hold can lead to flipping when it would otherwise not. Moreover, it can preclude flipping when flipping would otherwise be optimal.

Figure 12: Comparing Discrete and Continuous Labor Choice




Note: Labor supply of worker with $130,000 wage rate who is either restricted to full-time or part-time work (discrete labor) or not restricted (flexible labor). The left-hand side assumes our standard disutility of labor parameter, $\chi = 1$. The right-hand assumes $\chi = 1.15$. With flexible labor, the household flips around the FICA threshold for $\chi = 1$, yet stays below it for $\chi = 1.15$. With discrete-labor, the household always works full time for $\chi = 1$, yet sometimes works part time for $\chi = 1.15$.

Figure 12 considers a $130,000-wage worker who is limited to working either full or part time. Full time (half time) is defined as providing 1.2 (0.6) units of labor supply in a period. The left-hand panel is based on the same preferences as used above. The right-hand panel multiplies the disutility of work in each period by 1.15. The solid blue curve shows the optimal age-labor supply profile when labor supply is discrete. The dotted curve shows the corresponding fully flexible labor-supply case. In the left panel, restricting labor supply leads to no flipping when flipping around the FICA threshold would otherwise arise. In the right panel, the opposite occurs: Optimal discrete choice features flipping – occasionally working part time – when optimal flexible choice entails fixed annual labor supply. This has important implications for estimating structural labor-supply models. It suggests that assuming discrete choice to make one's model computationally tractable may be problematic.

## 6.2 Taxing Couples

Taxing couples on their joint income renders the two spouses' labor-supply decisions interdependent. To examine this inter-dependency, we assume that both household members are the same age and live for $T$ periods with joint lifetime utility given by

$$\sum_{t=1}^{T} \left(\frac{1}{1+\rho_t}\right)^{t-1} U(c_{1,t}, c_{2,t}, l_{1,t}, l_{2,t}), \tag{5}$$

where $c_i$ and $l_i$ are consumption and labor supply of household member $i = 1, 2$. Per-period utility satisfies:

$$U(c_1, c_2, l_1, l_2) = 2 \cdot \log\left(\frac{c_1 + c_2}{2}\right) - \frac{l_1^{1+1/\gamma}}{1+1/\gamma} - \frac{l_2^{1+1/\gamma}}{1+1/\gamma}. \tag{6}$$

Thus, spouses value average consumption, whereas their two disutilities of labor supply are simply added together. Both spouses retire at 65, both have the same time-preference rate, and both have the same Frisch elasticity of labor supply. Although positing couples facing joint taxation adds another labor-supply path over which the GLO must optimize, the program readily handles this. To be clear, we continue to randomly adjust two values of either consumption or labor supply in each iteration of our stochastic pattern-search routine. But now the two random adjustments are chosen from the set of annual labor supplies of each spouse and the annual levels of the two spouses' common consumption. Under *joint taxation*, household wage income is pooled and subject to the same marginal tax rates as in table 1, but with doubled levels of tax brackets. Under *separate taxation*, each spouse's wages are taxed separately according to table 1. The FICA tax treats spouses as single. As for the provision of basic income, couples whose joint labor earnings are less than \$30,000 receive \$20,000.

Figure 13 displays the labor supply of each spouse as well as taxable family labor income. Each row references a different couple with the higher earner's wage displayed in the first column and the lower earner's wage shown in the second column. The third column displays the couple's total taxable labor income. Blue lines reference the case of joint taxation and red lines denote separate taxation. The first row considers a relatively low-wage couple with wage rates of \$50,000 and \$30,000. There are two remarkable findings. First, under both tax schemes, the two household members coordinate to work less in some periods to collect the basic income. Second, when comparing the two tax schedules, joint taxation discourages the secondary earner from working, leaving that spouse working substantially fewer hours. This reduces overall taxable family income, even though joint taxation results in lower average taxes for any given pair of labor supply choices. This second observation also pertains to the case of our higher earning couple, with high and low spousal wage rates of $110,000 and $80,000, respectively. In this case, the primary earner flips their labor supply above and below the FICA threshold. The secondary earner, in contrast, works the same amount at all ages.

The third row considers a couple with even higher wage rates – $130,000 and $100,000. There is no longer any flipping: the primary earner remains above the FICA threshold while the secondary earner remains below. However, in this case, joint taxation not only lowers the secondary earner's labor supply. It also substantially increases the primary earner's labor supply. Consequently, taxable family income increases. In all three cases, the gap between the amount of labor supplied by the primary earner and the secondary earner increases in switching from separate to joint taxation. Joint taxation is always preferred by our married households. The welfare gain is substantial for the low-income case – roughly 5 percent as a CEV. It is marginal for the higher-earning couple – less than one percent.

## 6.3 Social Security

The GLO can also seamlessly handle Social Security's retirement-benefits formula, which, as noted, is a function of a worker's highest past 35 years of covered earnings. Including the system's benefit formula is simple. We just add it to our model as another fiscal program and let the GLO compute optimal labor supply, consumption, and savings paths incorporating the benefit formula's complex work incentives and disincentives.

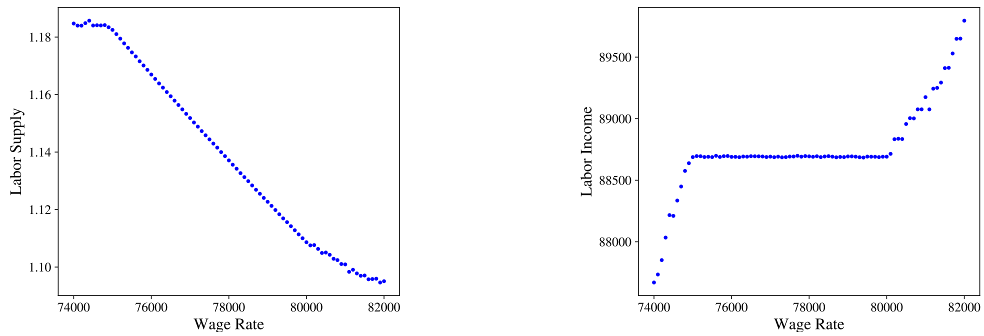
Generally, Social Security retirement benefits are determined in three steps. First, eligibility is established based on the accumulation of Social Security credits. A worker must earn at least 40 credits over their working life. In 2025, one credit corresponds to yearly gross earnings of $1,810, with a maximum of four credits obtainable per

Figure 13: Labor Supply when Couples are Taxed Jointly or Separately


Note: Labor supply and taxable family income over the life cycle for three different types of couples. In the first case, both earners coordinate to collect basic income in some periods. In the second case, the primary earner flips around the FICA threshold. In the third case the primary earner increases labor supply substantially when taxed jointly to make use of lower marginal rates.

year. Step two entails computing the worker's Average (Indexed) Monthly Earnings (AIME). This is done by selecting the 35 highest annual income values, summing them, and dividing by 420 (35 years × 12 months). Third, we calculate the Primary Insurance Amount (PIA), which forms the baseline retirement benefit. In 2025, the PIA is calculated as the sum of 90% of AIME up to $1,226, 32% between $1,226 and $7,391, and 15% of anything above $7,391. We then directly take the resulting annualized PIA as our benefit amount. Our stylized analysis departs from several Social Security benefit provisions. First, in the actual system, past earnings are indexed,

Figure 14: Bunching at upper PIA Bend Point

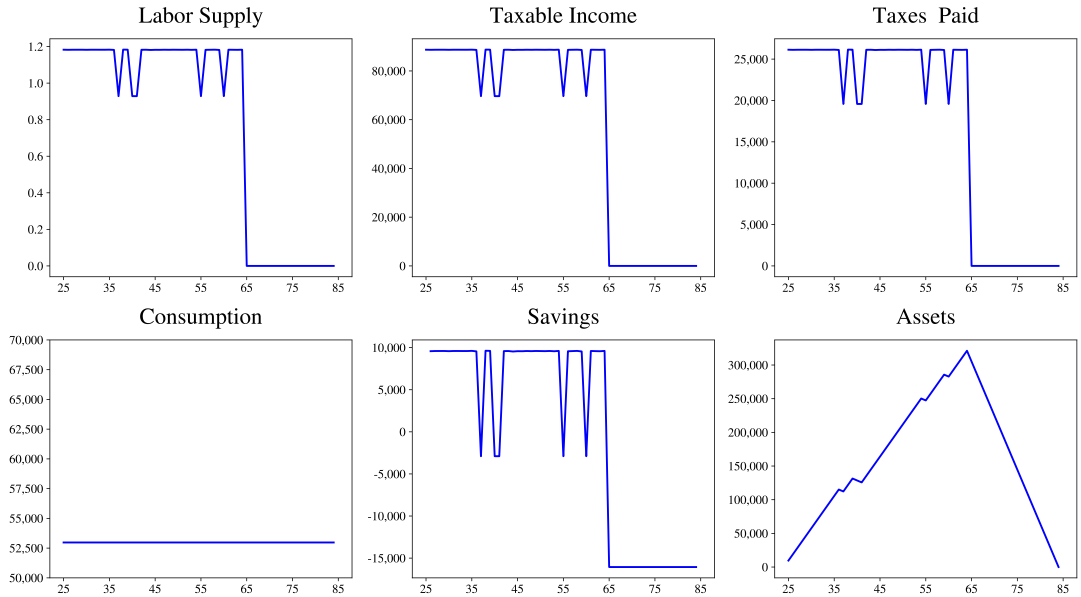



Note: Labor supply and labor income around the upper bend point of the PIA formula, located at \$7,391 monthly (i.e. \$88,692 yearly) income.

through age 60, to reflect growth in economy-wide average earnings. However, since our model assumes zero real nominal wage growth, it avoids this complication. Second, Social Security's retirement benefit depends on the age of first collection, with delayed collection providing a higher annual benefit. As our retirement age is fixed, we ignore both "Early Retirement Reductions" and "Delayed Retirement Credits" and set workers' retirement benefits equal to their PIAs. Third, U.S. retirement benefits are subject to federal and well as state-specific income taxation. Incorporating such taxation would introduce additional kinks, obscuring the analysis. For this reason, we omit benefit taxation.

Beyond showing that the GLO can readily handle this extended optimization problem, our results provide some interesting insights. For instance, figure 14 highlights how the sharp drop in the PIA benefit bracket at an AIME of \$7,391 induces substantial bunching around the corresponding annual gross earnings of \$88,692, as earning beyond that point produces far lower increases in the PIA and, thus, one's retirement benefit. Moreover, the fact that only the 35 highest yearly earnings matter for determining the PIA is reflected in the household's labor supply profile in figure 15. During 5 years, labor supply drops to a lower level, with the timing of these deviations being entirely random. In other words, compared to a hypothetical retirement system where all yearly income levels are considered identically in calculating retirement benefits, our household, which works for 40 years, increases its labor supply during 35 of those years and decreases it during 5 others, to exploit the asymmetric calculation of AIME.

Figure 15: Life-Cycle Profiles with Retirement Benefits




Note: Life-cycle profiles for a household with a wage rate of 75,000, receiving retirement benefits. As only the 35 highest yearly earnings are considered in determining retirement benefits, the household chooses to supply less labor during the remaining 5 years.

## 6.4 Optimal Choice under Life-Cycle Risk

Solving problems with NND budget sets that also include life-cycle risk poses a formidable challenge. One approach that could readily be implemented with our current version of the GLO is to follow Cai and Judd (2023) and assume certainty-equivalent behavior.[21]

The standard, but far more challenging path toward handling uncertainty requires the use of dynamic programming. A variant of the GLO might be used to find optimal policies in each step of value function iteration. While the GLO will likely be up to this task, a major additional challenge arises: interpolating the resulting non-differentiable value function accurately.[22]

[21] Agents, in this case, make decisions in each period as if all future shocks equal their expected values. This transforms lifetime uncertainty problems into sequences of deterministic problems, which, as we've seen, the GLO can readily handle. The concern here is whether this approach properly captures risk averse behavior, specifically precautionary consumption and labor supply.

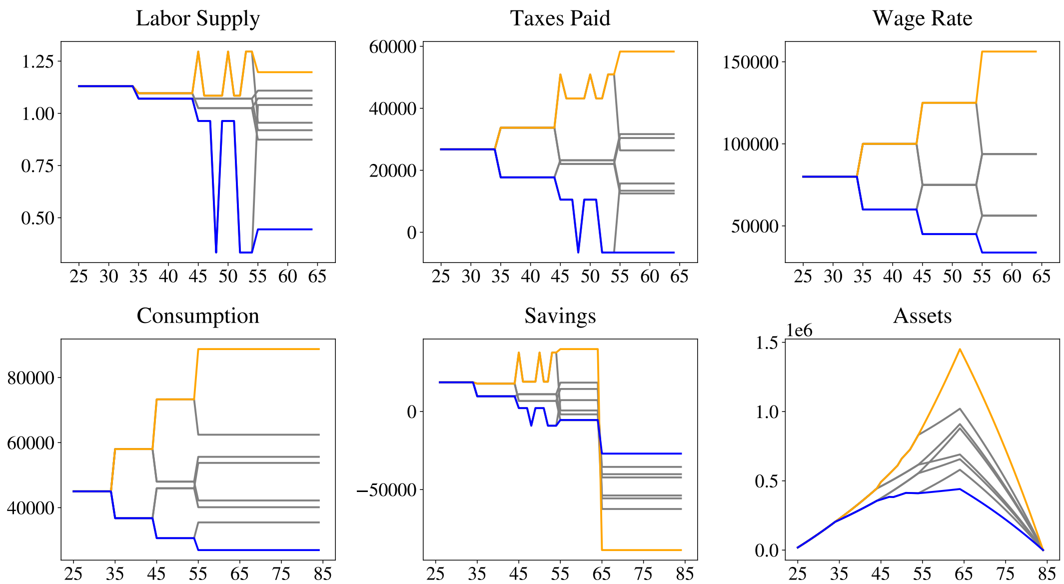
[22] A potential way to interpolate the value function in several dimensions while preserving monotonicity and concavity is Delaunay interpolation as employed in Brumm and Grill (2014).

A more straightforward extension of our approach is to use the GLO to maximize expected utility by simultaneously determining optimal decisions along all possible sequences of future random outcome paths. Conceptually speaking, this expanded program is fundamentally identical to the one that we've solved. However, it is clearly a far higher-dimensional problem. To maintain tractability, we consider wage-rate uncertainty, but limit the number of periods in which wage-rate shocks can occur. Our example nicely illustrates both the potential and limitations of the GLO when it comes to stochastic settings.

We assume life-cycle risk in the following form: every ten years, the household's wage-rate either increases or decreases by 25% with equal probability, as illustrated in the north-east corner of figure 16. For illustrative purposes we choose the initial wage rate so that the optimal solution covers all main characteristics of labor-supply responses to the tax code that we have discussed in section 5. The highlighted best (yellow) and worst (blue) paths in figure 16 are most interesting. First, the FICA threshold induces our household to flip if its wage rate increases at 35 and 45. Second, the basic income notch results in flipping in case the household's wage rate declines at 35 and 45. Third, that same notch generates bunching in case the household's wage rate falls further at 55.

The GLO performs without difficulty despite the added complexity. But its solution requires minutes instead of seconds, as the dimensionality of the problem increases from 100 in the deterministic case to 460 – labor supply and consumption choices across all contingencies. While the GLO in its current implementation is able to solve problems where shocks occur in up to ten periods, it is clear that solving life-cycle problems where shocks occur on a yearly frequency requires far greater computation capacity than appears to currently exist.

Figure 16: Life-Cycle Profiles under Risk




Note: Life-cycle and wage rate profiles for a household with an initial wage rate of $80,000. Random shocks to the wage rate occur every 10 years, resulting in either a 25% increase or decrease. The best (yellow) and worst (blue) scenario are highlighted.

# 7 Conclusion

This paper develops the Global Life-Cycle Optimizer to study the impacts typical elements of fiscal policy can have on work and saving decisions. The GLO is a stochastic pattern-search algorithm specifically designed to determine optimal economic behavior in the context of non-differentiable, non-convex, and discontinuous (NND) choice sets. To demonstrate the GLO's potential, we consider a simplified U.S. fiscal system comprising three elements – the federal personal income-tax brackets, the Social Security payroll tax with its taxable earnings ceiling, and the provision of basic income to workers earning below a threshold. This tax system comprises eight differently-sized kinks (seven convex and one concave) and one notch.

The GLO readily reproduces the anomalous behavior that theory predicts and data record – the bunching of earnings just below benefit-eligibility or higher marginal-tax thresholds. But it also produces flipping – switching in different years between over-time work (high labor supply) and part-time work (low labor supply). Of most surprise is the wide range of wages over which labor supply is dramatically altered.

This is particularly true for low-wage workers, some of whom reduce their labor supply by more than 50 percent in the course of bunching their wages.

As we show, the GLO can handle total income taxation, discrete choice, joint taxation of married couple's labor supplies, Social Security benefit-formula related work disincentives, and limited wage-rate uncertainty. Furthermore, the GLO can, it appears, simultaneously handle cash-flow constraints, fixed costs of working, minimum and maximum hours restrictions, labor supply adjustment costs, and benefit-program participation costs.[23] Inclusion of labor adjustment costs would surely limit the amount of flipping.[24] Agents who would otherwise flip absent adjustment costs would likely choose to work full time for part of their work span and part time for the remainder. This would manifest as early retirement.

The GLO also seems ideal for evaluating tax reforms. With a large and extensive data set and careful delineation of prevailing fiscal policies, one can compare GLO's results under current policy with those under proposed alternative policies. Although the GLO's capacities to handle the full US fiscal system remain to be seen, preliminary work with The Fiscal Analyzer, which encompasses the entire potpourri of U.S. federal and state policy (see Auerbach et al., 2023), suggests no difficulties. Intuitively, altering GLO's annual net tax function changes nothing fundamental.

A final point. One might reasonably ask whether individual households can make the calculations being modeled and processed by GLO. Complex fiscal provisions are, after all, a postwar phenomenon, not problems human brains have evolved to solve. Our response is twofold. First, observed earnings bunching shows that households can comprehend and appropriately respond to at least some forms of fiscal non-linearities. Second, over time, computational algorithms, like the GLO, will surely assist households in making optimal life-cycle labor supply and saving decisions.

[23] Moffitt (1983), Fraker and Moffitt (1988), and Moffitt (1992) incorporate decisions over plan participation.

[24] Adjustment costs may reflect search time needed to find jobs that permit desired hours of work. They can also reflect loss of firm-specific human capital, reduced accumulation of human capital, and depreciation of human capital.

# APPENDIX

## A Consumption Equivalent Variation

Denote lifetime utility by

$$W(c,l) = \sum_{t=1}^{T} \beta^{t-1} U(c_t, l_t), \text{ with } \beta = \left(\frac{1}{1+\rho}\right). \tag{7}$$

Let $(c^0, l^0)$ and $(c^1, l^1)$ be consumption and labor supply under the considered tax system and under lump-sum taxation, respectively. Then the consumption equivalent variation (CEV) is defined by the following equality:

$$W(c^0(1+CEV), l^0) = W(c^1, l^1).$$

For the left-hand side we get:

$$\begin{aligned}
W(c^0(1+CEV), l^0) &= \sum_{t=1}^{T} \beta^{t-1} U(c_t^0(1+CEV), l_t^0) \\
&= \sum_{t=1}^{T} \beta^{t-1} \left( \log[c_t^0(1+CEV)] - \chi \frac{(l_t^0)^{1+1/\gamma}}{1+1/\gamma} \right) \\
&= \sum_{t=1}^{T} \beta^{t-1} \left( \log(c_t^0) + \log(1+CEV) - \chi \frac{(l_t^0)^{1+1/\gamma}}{1+1/\gamma} \right) \\
&= \sum_{t=1}^{T} \beta^{t-1} \left( \log(c_t^0) - \chi \frac{(l_t^0)^{1+1/\gamma}}{1+1/\gamma} \right) + \sum_{t=1}^{T} \beta^{t-1} \log(1+CEV) \\
&= W(c^0, l^0) + \sum_{t=1}^{T} \beta^{t-1} \log(1+CEV).
\end{aligned}$$

Combining the above equations results in:

$$\begin{aligned}
&W(c^0(1+CEV), l^0) = W(c^1, l^1)\\
\Leftrightarrow\ &W(c^0, l^0) + \sum_{t=1}^{T} \beta^{t-1} \log(1+CEV) = W(c^1, l^1)\\
\Leftrightarrow\ &\sum_{t=1}^{T} \beta^{t-1} \log(1+CEV) = W(c^1, l^1) - W(c^0, l^0)\\
\Leftrightarrow\ &CEV = \exp\left(\frac{W(c^1, l^1) - W(c^0, l^0)}{\sum_{t=1}^{T} \beta^{t-1}}\right) - 1.
\end{aligned}$$

# B GLO Algorithm

This appendix provides a formal description of the working of GLO for a single starting guess. We avoid the outer loop for multiple starting guesses to simplify notation and as its implementation is trivial: a final step simply chooses the solution with the highest objective value among all solutions resulting from different starting guesses. For the above results, we ran the GLO with the following parameter values: $m^0 = 1$, $J = 500$, $b_l = 0$, $b_u = 1.25$, $\varepsilon = 10^{-8}$.

1. **Initialization:** Set stopping criterion $\varepsilon > 0$, initial mesh size $m^0 > \varepsilon$, and adjustment interval $[b_l, b_u]$. Set initial path for consumption and gross labor income of size $N = T + R$,[25]

   $$x^0 = \{c_1, \ldots, c_T, y_1, \ldots, y_R\}.$$

   Note that, for a given $x$, implied consumption is $c_t = x_t$, labor supply is $l_t = x_{T+t}/w_t$, and asset holdings are $a_{t+1} = w_t l_t + (1 + r_t)a_t - \mathcal{T}(w_t l_t) - c_t$ for $t = 1, \ldots, T$. Specify the objective function $W$ (with slight abuse of notation using gross labor income instead of labor supply as input) as follows:

   $$W(x) = \sum_{t=1}^{T} \left(\frac{1}{1+\rho_t}\right)^{t-1} U\left(x_t, \frac{x_{T+t}}{w_t}\right).$$

[25] We mostly use the following guess. A random labor supply fixed over working life, and consumption before (in) retirement as a random fixed fraction of labor income (in the last period before retirement), thus $x^0 = \{\tilde{c}w_1\tilde{l}, \ldots, \tilde{c}w_R\tilde{l}, \ldots, \tilde{c}w_R\tilde{l}, w_1\tilde{l}, \ldots, w_R\tilde{l}\}$, with $\tilde{c} \sim \mathcal{U}(0, 1), \tilde{l} \sim \mathcal{U}(0, 1.5)$.

2. **Construction of poll set:** Given the current point $x^k$ and mesh size $m^k$, the poll set contains $J$ points of dimension $N$,

$$\mathcal{P} = \{p_1, \dots, p_J\}.$$

To generate $p_j$, $j = 1, \dots, J$, execute the following steps:

   (a) Draw two distinct random integers $i_1, i_2 \in \{1, \dots, N\}$ and one uniformly distributed random real number $z \in [b_l, b_u]$.

   (b) Compute four different candidates $\tilde{p}_{j,1}, \dots, \tilde{p}_{j,4}$ that are identical to $x^k$ except for entries $i_1$ and $i_2$. In particular, set $\tilde{p}_{j,1} = \cdots = \tilde{p}_{j,4} = x^k$ and add and subtract $m^k$ to $i_1$, and $z \cdot m^k$ to $i_2$:

$$\begin{aligned}
\tilde{p}_{j,1}(i_1) &= x^k(i_1) + m^k, & \tilde{p}_{j,1}(i_2) &= x^k(i_2) + z \cdot m^k \\
\tilde{p}_{j,2}(i_1) &= x^k(i_1) + m^k, & \tilde{p}_{j,2}(i_2) &= x^k(i_2) - z \cdot m^k \\
\tilde{p}_{j,3}(i_1) &= x^k(i_1) - m^k, & \tilde{p}_{j,3}(i_2) &= x^k(i_2) + z \cdot m^k \\
\tilde{p}_{j,4}(i_1) &= x^k(i_1) - m^k, & \tilde{p}_{j,4}(i_2) &= x^k(i_2) - z \cdot m^k
\end{aligned}$$

   (c) For each $x = \tilde{p}_{j,1}, \dots, \tilde{p}_{j,4}$ and for $t = 1, \dots, T$, compute implied $a_{t+1}$. If $t < T$ and $a_{t+1} < 0$ or $t = T$, set $x_t = x_t + a_{t+1}$, ensuring that the borrowing constraint is just satisfied (for $t < T$) or all assets are consumed (for $t = T$).

   (d) Choose $p_j$ to be the candidate with the highest objective value:

$$p_j = \underset{\{\tilde{p}_{j,1},\dots,\tilde{p}_{j,4}\}}{\arg\max}\ W(p).$$

3. **Evaluation of poll set**: Pick the $p_i$ with the highest objective value:

$$p^* = \arg\max_{\mathcal{P}} W(p).$$

   If $W(p^*) > W(x^k)$, then the poll is successful: set $x^{k+1} = p^*$ and $m^{k+1} = m^k \cdot 2$. If $W(p^*) \leq W(x^k)$, the poll was unsuccessful: set $x^{k+1} = x^k$ and $m^{k+1} = m^k/2$.

4. **Convergence**: If $m^k < \varepsilon$, stop and return $x^{k+1}$ as solution; otherwise, go to step 2.

# C Global Optimization of Standard Test Functions

We follow Arnoud et al. (2019) and use standard test functions for global optimizers to benchmark the performance of GLO. In particular, we use the following three test functions:

- Levi function No. 13:

$$f(x) = \sin^2(3\pi x_1) + (x_n - 1)^2[1 + \sin^2(2\pi x_n)] + \sum_{i=1}^{n-1}(x_i - 1)^2[1 + \sin^2(3\pi x_{i+1})] + 1, \tag{8}$$

  with $x \in [-10, 10]^n$, which has its global minimum at $x = (1, \ldots, 1)$ with function value $f(1, \ldots, 1) = 1$.

- Rastrigin function:

$$f(x) = An + \sum_{i=1}^{n}[x_i^2 - A\cos(2\pi x_i)] + 1, \tag{9}$$

  with $A = 10$ and $x \in [-5.12, 5.12]^n$, which has its global minimum at $x = (0, \ldots, 0)$ with function value $f(0, \ldots, 0) = 1$.

- Griewank function:

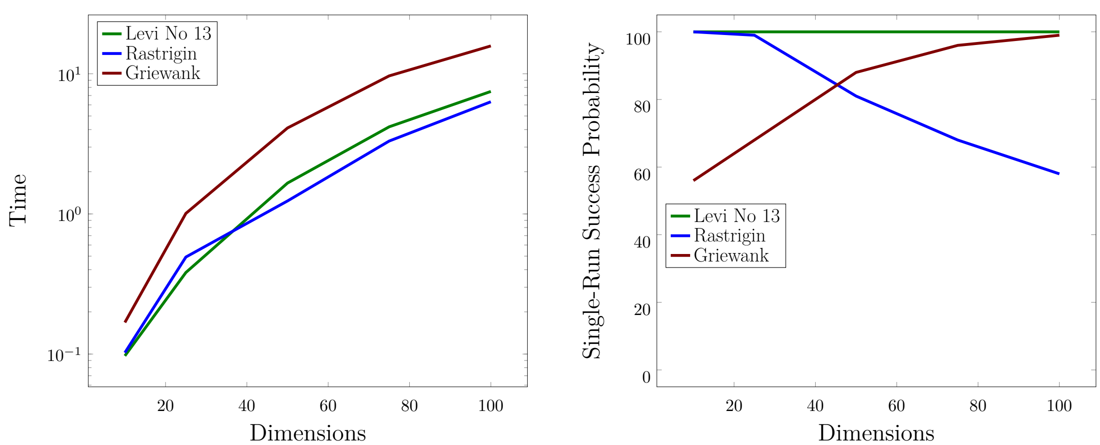
$$f(x) = \sum_{i=1}^{n}\frac{x_i^2}{a} - \prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right) + 2, \tag{10}$$

  with $a = 200$ and $x \in [-100, 100]^n$, which has its global minimum at $x = (0, \ldots, 0)$ with function value $f(0, \ldots, 0) = 1$.

Besides the change of objective function, we only make two obvious and two additional minor adjustments to the GLO. Clearly, we are now minimizing instead of maximizing and step 2 (c) is obsolete as we are now solving unconstrained optimization problems. Moreover, we choose $J = 2000$, i.e., we increase the size of the poll set in each iteration. This choice guarantees a success rate of at least 50% in all cases of the test functions as reported below. Finally, we set $\varepsilon = 10^{-6}$. This choice does not affect the success rate, as GLO either finds the global minimum or gets stuck in a local one, but it lowers the compute time to some degree.

Figure 17 displays the performance of the GLO on the three considered test functions of varying dimensionality. The left panel displays the compute time for 100 GLO runs, which turns out to grow much less than exponentially (concave shape with logarithmic time scale). The GLO can swiftly solve all three test functions in up to 100 dimensions.[26] It does so at a remarkable success rate of more than 50% across all test functions and dimensions considered, as the right panel of figure 17 shows the success rates of GLO to find the respective global minimum. We define a trial as successful if the absolute distance between the objective function value found by GLO and the global optimum is less than $10^{-6}$. While GLO finds the global optimum for the Levi No. 13 function in all trials, the success rate is increasing in the number of dimensions for the Griewank function and decreasing for the Rastrigin function. Note that these success rates of the GLO for single starting values translate to very high success rates for the GLO with multiple starting guesses. Even in the worst case, the 100-dimensional Rastrigin function, an implementation with 10 starting guesses has a success rate above 99.9%.

Figure 17: GLO's Performance on Standard Test Functions




Note: The left panel displays the average compute time (in seconds) for a single run of the GLO when solving the respective test functions. The right panel shows the success probability of the GLO for a single starting guess in finding the analytically known global minima of these test functions. With 10 starting guesses, success rates are above 99.9% for all considered specifications.

[26]Note that Arnoud et al. (2019) reports results for up to ten dimensional problems only. Note also that that paper uses a fourth test function, the Rosenbrock function. As the GLO can only solve that function for up to about 12 dimensions, we do not include the results here.

# D GLO versus Value Function Iteration

This appendix first describes the discrete value function iteration algorithm that we used to verify the GLO's solutions and to compare its performance. It then provides details of the hardware infrastructure used to run both the GLO and VFI.

We solve the NND life-cycle problem by VFI using beginning-of-period-assets $(1 + r_t)a_t$ as state variable and discretizing the choices labor supply, $l_t$, and assets, $a_{t+1}$. The algorithm automatically generates a grid of possible asset choices depending on the wage rate,[27] while for labor supply the lower and upper bounds are always set to 0 and 1.5. Overall, we work with 6,001 possible asset choices and 601 possible labor supply choices, representing a grid size of about 3,600,000 points. At a given age $t$ and for each point on the asset grid, we identify the optimal choices given the $t+1$-period value function, thereby generating the $t$-period value function. Finally, we obtain the VFI-solution of the NND life-cycle problem by reconstructing the optimal paths of labor supply, consumption, and assets from the obtained grids of optimal choices, starting from $t = 0$ and $a_0 = 0$.

To accelerate the algorithm, we made adjustments that limit the number of possible choices at each step to a realistic subset, thereby avoiding unnecessary computations. Specifically, we do not consider all possible asset choices from zero to the upper bound, but only those within a range of plus/minus the current wage rate from given starting assets, making necessary adjustments if this would include points off the grid. Furthermore, we distribute the task of finding the optimal choices for different starting assets in a given period $t$ across different cores.

We ran our programs on bwUniCluster 2.0. In each case, we utilized a single node of the cluster consisting of two sockets, each equipped with an Intel Xeon Gold 6230 processor with 40 cores and a frequency of 2.1 GHz, without any integrated accelerators. In terms of memory, we simply used the cluster's default setting of 1,125MB per CPU. To exploit these multiple cores, we used the multiprocessing package from the Python Standard Library. This allowed us to simultaneously run the GLO with different wage rates or different starting guesses, and to parallelize the search for the optimal points on the asset and labor supply grid when running VFI.

[27] We set the upper bound to 15 times the wage rate with special rules for very low wage rates to account for the effects of the basic income. Using \$0 as the lower bound, the grid is then constructed with a step size of 0.25% of the chosen wage rate.

## Declaration of Generative AI and AI-assisted Technologies in the Manuscript Preparation Process

During the preparation of this work the authors used *refine* in order to identify inconsistencies in notation, argumentation, and referencing. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

# References


Arnoud, A., Guvenen, F., and Kleineberg, T. (2019). Benchmarking global optimizers. NBER Working Paper 26340.

Audet, C. and Dennis, J. (2003). Analysis of generalized pattern searches. *SIAM Journal on Optimization*, 13(3):889–903.

Auerbach, A. J. (1985). The theory of excess burden and optimal taxation. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume 1, pages 61–127. Elsevier.

Auerbach, A. J. and Hines, J. R. (2002). Taxation and economic efficiency. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume 3, pages 1347–1421. Elsevier.

Auerbach, A. J., Kotlikoff, L. J., and Koehler, D. (2023). U.S. inequality and fiscal progressivity: An intragenerational accounting. *Journal of Political Economy*, 131(5):1249–1293.

Azinovic, M., Gaegauf, L., and Scheidegger, S. (2022). Deep equilibrium nets. *International Economic Review*, 63(4):1471–1525.

Bastani, S. and Selin, H. (2014). Bunching and non-bunching at kink points of the swedish tax schedule. *Journal of Public Economics*, 109:36–49.

Bertanha, M., McCallum, A. H., and Seegert, N. (2023). Better bunching, nicer notching. *Journal of Econometrics*, 237(2):105512.

Bick, A. and Fuchs-Schündeln, N. (2018). Taxation and labour supply of married couples across countries: A macroeconomic analysis. *The Review of Economic Studies*, 85(3):1543–1576.

Blomquist, S., Newey, W. K., Kumar, A., and Liang, C.-Y. (2021). On bunching and identification of the taxable income elasticity. *Journal of Political Economy*, 129(8):2320–2343.

Blundell, R. and Hoynes, H. W. (2004). Has' in-work'benefit reform helped the labor market? In *Seeking a Premier Economy: The Economic Effects of British Economic Reforms, 1980-2000*, pages 411–460. University of Chicago Press.

Blundell, R. and MaCurdy, T. (1999). Labor supply: A review of alternative approaches. In Ashenfelter, O. C. and Card, D., editors, *Handbook of Labor Economics*, volume 3, pages 1559–1695. Elsevier.

Blundell, R. and Shephard, A. (2012). Employment, hours of work and the optimal taxation of low-income families. *The Review of Economic Studies*, 79(2):481–510.

Brown, K. M. (2013). The link between pensions and retirement timing: Lessons from California teachers. *Journal of Public Economics*, 98:1–14.

Brumm, J. and Grill, M. (2014). Computing equilibria in dynamic models with occasionally binding constraints. *Journal of Economic Dynamics and Control*, 38:142–160.

Brumm, J. and Scheidegger, S. (2017). Using adaptive sparse grids to solve high-dimensional dynamic models. *Econometrica*, 85(5):1575–1612.

Burtless, G. (1976). Taxes, transfers and preferences for work among black married men. Mimeographed. Cambridge, Mass.: Massachusetts Inst. Tech.

Cai, Y. and Judd, K. L. (2023). A simple but powerful simulated certainty equivalent approximation method for dynamic stochastic problems. *Quantitative Economics*, 14(2):651–687.

Carroll, C. D. (2006). The method of endogenous gridpoints for solving dynamic stochastic optimization problems. *Economics letters*, 91(3):312–320.

Chetty, R., Friedman, J. N., Olsen, T., and Pistaferri, L. (2011a). Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from danish tax records. *The Quarterly Journal of Economics*, 126(2):749–804.

Chetty, R., Guren, A., Manoli, D., and Weber, A. (2011b). Are micro and macro labor supply elasticities consistent? a review of evidence on the intensive and extensive margins. *American Economic Review*, 101(3):471–475.

Danzinger, S. and Plotnick, R. (1981). Chapter 6: Income maintenance programs and the pursuit of income security. *The ANNALS of the American Academy of Political and Social Science*, 453(1):130–152.

Duarte, V., Fonseca, J., Goodman, A. S., and Parker, J. A. (2021). Simple allocation rules and optimal portfolio choice over the lifecycle. NBER Working Paper 29559.

Eklöf, M. and Sacklén, H. (2000). The hausman-macurdy controversy: Why do the results differ across studies? *Journal of Human Resources*, pages 204–220.

Fraker, T. and Moffitt, R. A. (1988). The effect of food stamps on labor supply: A bivariate selection model. *Journal of Public Economics*, 35(1):25–56.

Friedberg, L. (1998). The social security earnings test and labor supply of older men. *Tax policy and the economy*, 12:121–150.

Friedberg, L. (2000). The labor supply effects of the social security earnings test. *Review of Economics and Statistics*, 82(1):48–63.

Guner, N., Kaygusuz, R., and Ventura, G. (2012). Taxation and household labour supply. *The Review of Economic Studies*, 79(3):1113–1149.

Guvenen, F. (2011). Macroeconomics with heterogeneity: A practical guide. NBER Working Paper 17622.

Hagstrom, P. A. (1996). The food stamp participation and labor supply of married couples: An empirical analysis of joint decisions. *Journal of Human Resources*, pages 383–403.

Harberger, A. C. (1964). The measurement of waste. *American Economic Review*, 54(3):58–76.

Hausman, J. A. (1985). Taxes and labor supply. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of Public Economics*, volume 1, pages 213–263. Elsevier.

Hoynes, H. W. (1996). Work, welfare, and family structure: what have we learned? NBER Working Paper 5644.

Iskhakov, F., Jørgensen, T. H., Rust, J., and Schjerning, B. (2017). The endogenous grid method for discrete-continuous dynamic choice models with (or without) taste shocks. *Quantitative Economics*, 8(2):317–365.

Kaygusuz, R. (2010). Taxes and female labor supply. *Review of Economic Dynamics*, 13(4):725–741.

Keane, M. and Moffitt, R. A. (1998). A structural model of multiple welfare program participation and labor supply. *International Economic Review*, 39(3):553–589.

King, R. G., Plosser, C. I., and Rebelo, S. T. (1988). Production, growth and business cycles: I. the basic neoclassical model. *Journal of Monetary Economics*, 21(2-3):195–232.

Kotlikoff, L. J. (1978). Social security, time for reform. In Boskin, M. J., editor, *Federal Tax Reform: Myth or Reality*, pages 119–144. Institute for Contemporary Studies.

MaCurdy, T., Green, D., and Paarsch, H. (1990). Assessing empirical approaches for analyzing taxes and labor supply. *Journal of Human Resources*, 25(3):415–490.

Maliar, L., Maliar, S., and Winant, P. (2021). Deep learning for solving dynamic economic models. *Journal of Monetary Economics*, 122:76–101.

Meyer, B. D. and Rosenbaum, D. T. (2001). Welfare, the earned income tax credit, and the labor supply of single mothers. *The Quarterly Journal of Economics*, 116(3):1063–1114.

Moffitt, R. A. (1983). An economic model of welfare stigma. *The American Economic Review*, 73(5):1023–1035.

Moffitt, R. A. (1992). Incentive effects of the us welfare system: A review. *Journal of Economic Literature*, 30(1):1–61.

Moffitt, R. A. (2002). Welfare programs and labor supply. In Auerbach, A. J. and Feldstein, M., editors, *Handbook of public economics*, volume 4, pages 2393–2430.

Moffitt, R. A. (2003). The temporary assistance for needy families program. In Moffitt, R. A., editor, *Means-tested transfer programs in the United States*, pages 291–364. University of Chicago Press.

Moore, R. and Pecoraro, B. (2020). Macroeconomic implications of modeling the internal revenue code in a heterogeneous-agent framework. *Economic Modelling*, 87:72–91.

Moore, R. and Pecoraro, B. (2021). A tale of two bases: Progressive taxation of capital and labor income. *Public Finance Review*, 49(3):335–391.

Moore, R. and Pecoraro, B. (2023). Quantitative analysis of a wealth tax for the united states: Exclusions and expenditures. *Journal of Macroeconomics*, 78:103559.

Pencavel, J. (1986). Labor supply of men: a survey. In Ashenfelter, O. C. and Layard, R., editors, *Handbook of Labor Economics*, volume 1, pages 3–102. Elsevier.

Rust, J. P. (1989). A dynamic programming model of retirement behavior. In Wise, D. A., editor, *The Economics of Aging*, pages 359–404. University of Chicago Press.

Rust, J. P. (1990). Behavior of male workers at the end of the life cycle: An empirical analysis of states and controls. In *Issues in the Economics of Aging*, pages 317–382. University of Chicago Press, 1990.

Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, 2(3):180–212.

Torczon, V. (1997). On the convergence of pattern search algorithms. *SIAM Journal on Optimization*, 7(1):1–25.

Zabalza, A., Pissarides, C., and Barton, M. (1980). Social security and the choice between full-time work, part-time work and retirement. *Journal of Public Economics*, 14(2):245–276.