Income Taxation with Frictional Labor Supply

Nicolas Werquin*

December 6, 2015

Abstract

This paper characterizes the optimal labor income taxes in an environment where individual labor supply choices are subject to frictions. Agents incur a fixed cost of adjusting their hours of work in response to changes in their idiosyncratic wages or the tax rates. This fixed cost can be thought of as the cost of searching for a new job in an economy where hours are constrained within the firm. I derive a formula that characterizes the optimal long-run progressive tax schedule in this economy. Adjustment frictions generate endogenously an extensive margin of labor supply conditional on participation. In addition to the standard intensive margin disincentive effects of taxes, the optimal tax schedule takes into account their effects on the endogenous option value of adjusting hours of work, and therefore depends on several new elasticities and marginal social welfare weights. I evaluate the quantitative magnitude of these novel theoretical effects and show that for a given intensive margin labor supply elasticity, the optimal long-run tax schedule is less progressive than a frictionless model would predict. The welfare miscalculations by wrongly assuming that labor supply is frictionless can be large, and are decreasing in the size of the intensive margin labor income elasticity.

*Toulouse School of Economics. Email: nwerquin@gmail.com. Download the latest version using this link. I am extremely grateful to Mikhail Golosov, Giuseppe Moscarini and Aleh Tsyvinski for their invaluable guidance and support. I thank Andrew Abel, Costas Arkolakis, Andrew Atkeson, Alan Auerbach, Martin Beraja, David Childers, Tibor Heumann, Johannes Hörner, Francesco Lippi, Costas Meghir, Kumar Muthuraman, Tommaso Porzio, Ali Shourideh, Alexis Toda, and Danny Yagan, for helpful comments and suggestions.
1 Introduction

Theoretical models of optimal labor income taxation typically assume that labor supply can be adjusted costlessly and optimally in response to wage or tax changes. The individual labor income elasticity, which measures the response of the worker’s taxable labor income to a percentage change in the net-of-tax rate, is then the crucial parameter governing the effects of raising taxes on government revenue. The optimal tax rate is determined by the trade-off between the revenue effect of raising taxes (including these behavioral losses) and the welfare effect (expressed in terms of public funds), measured by the marginal utility of consumption, or marginal social welfare weights. Even models of taxation that incorporate explicitly an extensive margin of labor supply typically consider the binary decision of whether to participate in the labor force, keeping the assumption that conditional on participation hours are either exogenously fixed or fully flexible on the intensive margin. A vast theoretical and empirical literature is thus devoted to the analysis and estimation of the labor supply elasticity, or taxable income elasticity, with respect to the marginal tax rates.

A common criticism of this framework for analyzing taxation is that the modeling of individual labor supply is unrealistic. Indeed, a large and growing body of empirical evidence shows that the adjustment of labor supply in response to productivity, wage or tax changes, is subject to substantial frictions. Specifically, workers in a given job face hour constraints set by firms and pay search costs to find new jobs. In other words, they must change jobs in order to adjust their hours of work, and that entails large fixed costs. The presence of such fixed adjustment costs generates endogenously an extensive margin of labor supply, conditional on participation, where the thresholds (i.e., the timing and the size) of adjustment are chosen optimally by the worker. There is little theoretical work that explicitly incorporates such adjustment frictions into models of labor income taxation to describe the social welfare effects of taxes in this more realistic context. Several questions arise, some of which have been partially studied in the literature: the sluggish response of individuals to tax changes may generate large long-run labor income elasticities for small short-run (observed) elasticities, as well as non-trivial welfare effects in the short-run if taxes increase at a time where individuals are not maximizing their instantaneous utility. The question I address in this paper instead is whether, in the presence of labor supply adjustment frictions in a dynamic economy, long-run optimal taxes differ from those implied by the frictionless, intensive margin, models, and if so, what are the theoretical forces that determine optimal policy in the frictional economy. The reason for studying this question is that one often (loosely) argues that the static frictionless model is aimed at capturing the long-run effects of taxes, once individuals have had the time to fully adjust their labor supply in response to the new tax rates; in other words, there may exist adjustment frictions at the individual level, but they become irrelevant in the long-run. I show in this paper that this reasoning is not correct when labor supply adjustments to wage and tax changes are optimally chosen and the government has access to non-linear tax instruments. In this case, non-linear taxes interact with the fixed costs at the micro level to generate long-run real effects of policy. The reason is that in addition to their standard intensive margin disincentive effects, non-linear taxes also affect the (endogenous) option value of adjusting behavior in response to wage changes, which
in turn impacts welfare. These effects do not disappear in the long-run, because individuals still face idiosyncratic shocks even though the economy is in a stationary steady-state. And micro frictions do not wash out in the aggregate when the policy tools that are available are non-linear.

I set up an analytically tractable dynamic (continuous-time) model in which individuals choose their labor supply as a function of their stochastic idiosyncratic wage shocks and the non-linear tax schedule. The individual wage is exogenous and follows a random growth process with jumps, consistent with both micro and macro evidence. The tax schedule is restricted to have a constant rate of progressivity, a specific functional form that closely approximates the actual U.S. tax and transfer schedule. To keep the model tractable I assume that individuals are born (or enter the labor force) and die (or retire) at an exogenous Poisson rate, and that they cannot save or borrow. In order to adjust their hours of work in response to wage or tax changes, individuals must pay a fixed cost. This fixed adjustment cost can be thought of as the cost of searching for a new job in an economy where hours are constrained within the firm, and is assumed proportional to the worker’s foregone utility of consumption from the search activity. Once they decide to pay the fixed cost, i.e., to start searching for a new job, they receive a job offer, or a costless adjustment opportunity, at an exogenous Poisson rate, which captures in a reduced-form way the frictions on the demand side of the labor market. As a result, hours of work evolve in a lumpy manner at the individual level: workers remain inactive, that is keep the same job, until their productivity (or wage) is such that their optimal (frictionless) labor supply is far enough from their current, actual labor supply; at this point they pay the fixed cost and start searching for a new job.\footnote{For salaried workers, an alternative setup would place the fixed cost on income rather than hours of work. All the results of this paper would continue to hold in this model, see Werquin (2014) for details.} I show that the optimal range of inaction is an interval that I characterize analytically, and that the aggregate stationary income distributions have Pareto tails that can be written transparently in closed form as a function of the progressivity of the tax schedule.

I then derive formulas characterizing the optimal long-run progressive tax schedule in this economy, that is, the tax rates that maximize utilitarian social welfare subject to a government budget constraint. In the frictionless model, I show that the optimum is characterized by sufficient-statistic expressions extending the standard static formulas to the steady-state of the dynamic model. These formulas are written in terms of the individual intensive margin labor income elasticities and marginal social welfare weights that I define, capturing respectively the revenue and welfare effects of perturbing the tax rates. In the frictional model, I first show that the effect of a uniform increase in the marginal tax rates is given by the same formula as in the frictionless setting. In particular, the relevant labor income elasticity entering this formula is the individual elasticity of frictionless income, even though in the presence of frictions the individual’s actual hours are generally different from her frictionless optimum. Intuitively, in the long run all individuals have had time to adjust their behavior to the new tax system and frictions wash out in the aggregate. I then turn to the effects of an increase in the progressivity of the tax schedule. There are several new effects that appear in the frictional economy that are not captured by the frictionless optimal taxation formula. First and most important, in the frictional economy the individual’s option value
of adjusting labor supply is endogenous to taxes. This is because an increase in progressivity reduces the volatility of the income process, as higher incomes are taxed at a higher rate, which in turn reduces the option value and narrows the inaction region. I show that this induces non-zero effects on revenue and welfare unless these two forces exactly cancel out. I define and characterize novel and intuitive, “extensive margin” elasticities and marginal social welfare weights that summarize these effects in the optimal tax formula. Moreover, the presence of adjustment frictions implies that individuals who earn the same income differ in their utility, as the least productive of them (i.e., those with a lower wage) are working more hours to earn the same income; this non-degenerate distribution is itself endogenous to the tax schedule. By treating the population earning a given income as a representative agent instead, the frictionless model thus ignores this endogenous heterogeneity and miscalculates the welfare effects of perturbing taxes. This effect is captured by another novel marginal social welfare weight in the optimal tax formula. Therefore, theoretically, the extensive margin of labor supply endogenously generated by hour requirements within jobs has non-trivial long-run effects on tax revenue and social welfare beyond the standard intensive margin effects.

Finally I calibrate the model and estimate numerically the magnitude of these novel theoretical effects. The first finding is that the extensive margin elasticities with respect to the rate of progressivity are non-negligible - of the same order of magnitude as the participation elasticities found in the literature. In the baseline calibration, however, these elasticities induce only a small behavioral effect on tax revenue: the revenue losses from raising progressivity calculated by wrongly assuming a frictionless economy are about 2% away from their true values. The reason is that the extensive margin elasticities are bounded, while an increase in progressivity induces unboundedly large changes in the marginal tax rates as income grows, and hence much larger intensive margin income responses to taxes. However, the welfare effects generated by the endogenous option value of adjusting labor supply and by the endogenous heterogeneity within incomes are larger, and imply benefits of raising the progressivity of the tax schedule relative to the frictionless model: the frictionless values are more than 7% away from their true value, leading to a lower optimal rate of progressivity. In contrast, assuming a larger intensive margin labor income elasticity (ε = 1 rather than ε = 0.33) dwarves these effects, as the standard intensive margin terms then dominate the new extensive margin terms; in this case, the frictionless model closely approximates the true long-run optimal tax rates. The reason why the extensive margin effects on welfare tend to reduce the gains of raising progressivity is the following. I explained above that the standard marginal social welfare weights capture the true welfare effects of taxes if the decrease in the volatility of incomes due to an increase in progressivity is exactly compensated by the narrowing of the inaction region (option value effect). In general, however, the latter effect is dominated by the former, so that an increase in progressivity is equivalent to a wider dispersion of individual incomes around their desired frictionless values, which adversely affects welfare, relative to the frictionless benchmark.

Finally, the results I derive in this paper apply more broadly than to this taxation theory framework. The key finding of this paper is that non-linear policies interact with fixed adjustment costs to yield long-run aggregate real effects on both efficiency and welfare. This is in contrast with a large literature that studies models with fixed costs at the micro level, where typically the
aggregate economy behaves in the long-run as a frictionless (representative agent) model. I show in this paper that this insight is correct only if the available policy instruments are linear and thus do not affect the option value of adjusting behavior. This insight should apply generally to models with fixed costs and non-linear policy tools.

**Related literature.** A crucial variable in this paper is the individual (Hicksian) labor income elasticity to the marginal tax rates. The empirical literature estimating these elasticities is vast: see, e.g., Saez, Slemrod, and Giertz (2012) and Keane and Rogerson (2012) for recent surveys. Importantly, there is a large empirical literature that points to the presence of frictions in the adjustment of labor supply. Altonji and Paxson (1992) show that changes in labor supply preferences have a much larger effect on hours of work when individual change jobs, suggesting that adjusting behavior entails substantial fixed costs. Other papers have similarly argued that labor supply is constrained by adjustment costs and hours requirements set by firms, e.g., Cogan (1981), Altonji and Paxson (1988), Dickens and Lundberg (1993), Chetty, Friedman, Olsen, and Pistaferri (2011), Gelber, Jones, and Sacks (2013). Chetty, Guren, Manoli, and Weber (2012) and Chetty (2012) argue that adjustment frictions can fully explain the wide range of estimates of the steady-state elasticities found in the literature. Holmlund and Söderström (2008) argue that the short-run and the long-run elasticities may differ, also consistent with adjustment costs. My contribution is to incorporate explicitly these fixed costs into a dynamic taxation framework and derive the consequences for long-run optimal income taxes. These fixed costs generate endogenous extensive margin responses for employed individuals and long-run elasticities that differ from the short-run elasticities due to the sluggish adjustment of hours.

This paper also relates to several strands in the optimal taxation literature. My frictionless model is related to that of Heathcote, Storesletten, and Violante (2014), who also restrict the set of available tax instruments to two-parameter schedules, and analyze the effects of progressivity on social welfare with imperfect private insurance and investment in skills. My frictionless model is simpler than theirs, which allows me to tractably introduce and study the implications of frictional labor supply. The literature on the sufficient statistic approach to taxation, see e.g. Saez (2001), Chetty (2009), Golosov, Tsyvinski and Werquin (2014), derives optimal tax formulas for a very large class of models, irrespective of the underlying functional forms for the utility functions, the sources of heterogeneity, etc. However, these models generally implicitly assume that labor supply can always be set optimally at no cost. In this paper I show that these sufficient statistic formulas do not generally hold, even in the long run, in the presence of simple adjustment frictions. There is a theoretical taxation literature with labor supply responses on the extensive margin. Saez (2002), Choné and Laroque (2011), Jacquet, Lehmann, and Van der Linden (2013), Shourideh and Troshkin (2014), Lehmann, Kroft, Kucko and Schmieder (2015) study optimal taxation problems where individuals face a fixed cost of working, leading to binary participation decisions. Rogerson and Wallenius (2009) and Ljungqvist and Sargent (2001) study the labor supply elasticities in numerical models with intensive and extensive margins. I extend these papers' insights by generating and studying more general sources of extensive margin responses of labor supply to taxes, namely,
conditional on participation. It would be straightforward to include an explicit participation margin in my setting (similar to Alvarez, Borovickova and Shimer (2015)), although empirically much of the difference in labor supply across countries with different tax regimes is driven by hours worked conditional on employment (see Davis and Henrekson (2005), and Chetty, Guren, Manoli and Weber (2011)). Chetty, Looney, and Kroft (2009) propose a model of bounded rationality (where the fixed adjustment cost is interpreted as a cognitive cost) where individuals’ responses to taxes are affected by tax salience, and show that this feature affects the calculation of the impact of taxes on social welfare. Chetty, Friedman, Olsen, and Pistaferri (2011) study a model in which labor supply is subject to search costs and jobs are characterized by hours requirements. These models are primarily static, and do not capture the dynamic decisions of individuals based on their option value of waiting to adjust, nor the long-run effects of non-linear taxes on social welfare. A paper related to mine is that of Alvarez, Borovickova and Shimer (2015) who also model labor supply decisions as a stopping time problem. They only consider the transitions between employment and unemployment, however, and do not focus on the implications of this class of models for optimal taxation.

Finally, the technical tools I use to analyze my model of individual behavior are those developed in the impulse control literature originally developed to analyze operations research questions. Dixit and Pindyck (1994) and Stokey (2008) summarize many references and applications of these models to economics, primarily on monetary and investment topics. Richards (1977), Harrison, Sellke and Taylor (1985), Bertola and Caballero (1994), Grossman and Laroque (1990), Caballero and Engel (1999), and more recently Alvarez and Lippi (2013) and Alvarez, Le Bihan, and Lippi (2014), to cite only a few, have made important theoretical contributions to this literature, on which this paper builds. In public finance, there is a rich literature on investment in the presence of adjustment costs: Hall and Jorgensen (1967), Summers (1981), Abel (1983), Auerbach and Hines (1987), Auerbach (1989), Auerbach Hassett (1992). I bring this literature to the study of labor supply, since we now know that labor adjustment costs and not just capital adjustment costs can be important, and in addition I study optimal policy in this class of models.

The structure of the paper is as follows. Section 2 sets up the environment and describes the maximization problems of the individual and the government. Section 3 analyzes the optimal individual behavior. Section 4 characterizes the aggregate steady-state of the economy. Section 5 derives formulas for optimal taxes. Section 6 contains the calibration of the model and the numerical exercises. Section 7 concludes. The proofs of all the results are gathered in the Appendix.

2 Environment

There is a continuum of mass one of individuals in the economy. Time is continuous.

Preferences. Individuals have the following Greenwood, Hercowitz and Huffman (1988) utility function of consumption $c$ and hours of work ($labor$ $supply$) $h$, with isoelastic disutility of labor
supply:

\[ U(c, h) = \frac{1}{1 - \gamma} \left( c - \frac{1}{1 + 1/\varepsilon} h^{1+1/\varepsilon} \right)^{1-\gamma}, \]  

with \( \gamma \in [0, 1) \). They discount the future at rate \( \rho_1 \). They are born (or enter the labor force) and die (or retire) at an exogenous and constant Poisson rate \( \rho_2 \). I denote \( g(x) = (1 - \gamma)^{-1} x^{1-\gamma} \).

**Technology.** Individual productivity \( \theta \) is exogenous. The production function is linear in the labor input, so that in equilibrium workers’ wages \( w_t \) are always equal to their marginal productivity \( \theta_t \), and they can freely choose to provide any amount of labor supply \( h(\theta_t) \). Therefore, in the sequel, I substitute each worker’s exogenous productivity \( \theta \) with her wage \( w \).

An individual’s wage (i.e., productivity) at birth, \( w_0 \), is drawn from a log-normal distribution with mean \( m_w \) and variance \( s_w^2 \), i.e., \( f_{w_0}(\cdot) \sim \log\mathcal{N}(m_w, s_w^2) \). The idiosyncratic wage \( w_t \) then evolves stochastically over time \( t \geq 0 \) according to a geometric Brownian motion with expected growth rate \( \mu_w + \frac{1}{2} \sigma_w^2 \) and volatility \( \sigma_w \):

\[ d \ln w_t = \mu_w dt + \sigma_w dW_t, \]  

where \( W_t \) is a Wiener process, which generates a filtration \( \{\mathcal{F}_t\}_{t \geq 0} \). Individuals know their wage process (2) and observe its realizations at every instant \( t \).

The reduced-form equation (2) for the exogenous wage process can be microfounded, see e.g. Gabaix, Lasry, Lions and Moll (2015), and the references therein. This random growth process forms the basis of a leading theory of income inequality, as it naturally generates Pareto tails for the wage distributions (see, e.g., Gabaix (2009) and Section 4 below), a stylized fact that moreover plays an important role in the optimal taxation literature (Saez (2001)). The empirical literature (see Meghir and Pistaferri (2011) for a survey) estimates wage specifications of this form and its findings are consistent with the presence of a unit root in the wage process \( w_t \), that is, permanent wage shocks.

None of the results of this paper would be affected if wages were allowed to jump,\(^2\) i.e., \( d \ln w_t = \mu_w dt + \sigma_w dW_t + \nu_{w,t} dJ_t \), where \( dJ_t \) is a jump process with intensity \( \iota \) and innovations \( \nu_{w,t} \) are drawn from a given exogenous distribution \( f_{\nu} \). In this case there is a jump in \( [t, t + dt) \) (i.e., \( dJ_t = 1 \)) with probability \( \iota dt \) and no jump (i.e., \( dJ_t = 0 \)) with probability \( 1 - \iota dt \). Assuming a double-Pareto distribution \( f_{\nu} \), the jump process would imply that the distribution of income growth rates \( d \ln w_t \) itself also has Pareto tails, consistent with the evidence presented in Guvenen, Karahan, Ozkan and Song (2014). I ignore such jumps in the sequel to simplify the theoretical exposition.

**Budget constraint and taxes.** An individual with wage \( w \) who works \( h \) hours earns taxable labor income \( y = w \times h \) and pays taxes \( T(y) \) to the government. I assume that she cannot save or

---

\(^2\)In particular, see Chapter 6 in Oksendal and Sulem (2007) for an analysis of impulse control models with jump diffusions.
borrow, so that she consumes her net income at every instant:

\[ c = y - T(y). \]

The tax-and-transfer system is restricted within a class of two-parameter schedules (see, e.g., Benabou (2002), Heathcote, Storesletten, and Violante (2014)), defined as

\[ T(y) = y - \frac{1 - \tau}{1 - p} y^{1-p}, \]

with \( \tau \in \mathbb{R} \) and \( p < 1 \). I denote the tax schedule interchangeably by \( T(\cdot) \) or \((\tau, p)\). The parameter \( p \) is the coefficient of marginal rate progression (see Musgrave and Thin (1948)), or progressivity of the tax schedule. It is equal to the elasticity of the net-of-tax rate with respect to taxable income,

\[ p = - \frac{d \ln (1 - T'(y))}{d \ln y}. \]

If \( p = 0 \), the income tax schedule is linear with constant marginal tax rate \( \tau \). If \( p \in (0, 1) \), the ratio of the marginal tax rate to the average tax rate is \( T'(y) / [T(y) / y] > 1 \), so that the tax schedule is progressive. If \( p < 0 \), the tax schedule is regressive. Note that the marginal and average tax rates are monotone in earnings, and that average tax rates are negative for incomes \( y \) below \( \left( \frac{1 - \tau}{1 - p} \right)^{1/p} \).

This functional form for the tax schedule provides a close approximation of the actual tax system in the U.S. (see Heathcote, Storesletten, and Violante (2014)). The first panel of Figure 1 shows the marginal and average tax rates for two values of the progressivity parameter: \( p = 0.151 \), which is calibrated to the rate of progressivity of the US tax code (see Section 6) and \( p = 0.156 \). The second panel graphs these tax schedules at the bottom of the income distribution.

**Figure 1: Tax schedule**
Individual problem. Individuals choose their labor supply $h_t$ endogenously as a function of their information at time $t$. Consider a frictionless environment first, where they can adjust their labor supply optimally and costlessly at every instant. Denote by $V^* (w_0)$ (or equivalently $V^* (y_0)$) the value function of a worker with current wage $w_0$ (or income $y_0$). She solves the following problem:

$$V^* (w_0) = \max_{(h_t)_{t \geq 0}} \mathbb{E}_0 \left[ \int_0^\infty e^{-(\rho_1 + \rho_2) t} U (w_t h_t - T (w_t h_t), h_t) \, dt \right],$$

subject to the law of motion (2) of the wage. The solution to this problem gives the agent’s frictionless, or desired, labor supply $\{h^*_t\}_{t \geq 0}$ and consumption $\{c^*_t\}_{t \geq 0}$.

I now suppose that in order to adjust her labor supply (which I also refer to as a “job”) from $h$ to $h'$, the individual must pay a fixed (utility) cost $\kappa \geq 0$, which can be interpreted as the search cost of finding a new job (see the empirical evidence presented in, e.g., Altonji and Paxson (1992) and Chetty, Friedman, Olsen and Pistaferri (2011)). Formally, I assume that $\kappa$ is proportional to the utility $g(c^*)$ from the foregone (frictionless) disposable income due to the search activity: \(^3\)

$$\kappa = \kappa \times g(c^*),$$

with $\kappa \geq 0$.

After paying the fixed adjustment cost, the individual waits until she receives a job offer. Offers arrive at an exogenous Poisson rate $q$. Whenever she receives one, she can adjust her hours $h$ optimally and costlessly given her current productivity (i.e., wage) $w$. As long as she does not receive the offer, she keeps the same labor supply. \(^4\) Intuitively, $q$ captures in a reduced-form way the frictions on the demand side of the labor market: the larger the $q$, the faster an individual searching for a job finds one. \(^5\) As $q \to \infty$ with $\kappa > 0$, the model converges to a two-sided $(S, s)$ model similar to those studied in the operations research, monetary or investment literatures (see, e.g., Richards (1977), Harrison, Sellke and Taylor (1985), Dixit and Pindyck (1994), Stokey (2004), and the references therein). In this case the hours adjustments are entirely driven by labor supply considerations, i.e., productivity, taxes, and search costs. As $\kappa \to 0$ with $q < \infty$, on the other hand, the model converges to an environment similar to that of Calvo (1983); in this case adjustments are driven purely by the exogenous labor demand.

Individuals decide when and by how much to adjust their labor supply as their wage evolves. They can choose their hours optimally and costlessly at birth. We define an impulse control policy $p$.

---

\(^3\)Assuming $\gamma < 1$ in (1) ensures that the fixed cost $\kappa$ is strictly positive and increasing in income.

\(^4\)In a previous version of this paper (Chapter 1 in Werquin (2015)), I considered salaried workers instead, and assumed that the total taxable income $y$, rather than hours $h$, is subject to the fixed adjustment cost (see, e.g., the empirical evidence presented in Gelber, Jones and Sacks (2013)). In this case, $w$ and $h$ are interpreted respectively as productivity and effort, and their product $y$ is the agent’s effective labor supply, or income. As the individual becomes more productive and stays in her current job ($w$ increases and $y$ remains constant), she needs to provide less effort ($h$ decreases) to produce the required amount $y$. She adjusts her income upwards (resp., downwards) when she becomes so productive (resp., unproductive) that she spends most of her time idle (resp., when she must provide too much effort) to produce $y$. The results of the paper are unaffected by this alternative specification.

\(^5\)In this paper I focus on the labor supply effects of taxation, and thus assume that the labor demand frictions, summarized by the parameter $q$, are exogenous to taxes.
as a sequence of stopping times (the timing of the payments of the adjustment cost) \(0 < \tau_1 < \ldots < \tau_i < \ldots\) adapted to \(\{\mathcal{F}_t\}\), and a sequence of random variables (the size of log-hours adjustments upon receiving an adjustment offer) \(\Delta_0, \Delta_1, \ldots, \Delta_i, \ldots\), constructed inductively as follows. Let \(\tau_0 = \hat{\tau}_0 = 0\) and \(\{\hat{\tau}_i\}_{i \geq 1}\) be a sequence of i.i.d. random variables (waiting times to receive an adjustment offer after payment of the fixed cost) with exponential distribution \(\mathcal{E}(q)\), i.e., \(\mathbb{P}(\hat{\tau}_i \geq t) = e^{-qt}\) for all \(t \geq 0\). For all \(i \geq 0\), let \(\Delta_i\) be measurable with respect to the minimum \(\sigma\)-algebra \(\mathcal{F}_{\tau_i+\hat{\tau}_i}\) of events up to time \(\tau_i + \hat{\tau}_i\), and let \(\tau_{i+1} > \tau_i + \hat{\tau}_i\). I denote by \(\mathcal{P}\) the set of such impulse control policies \(p\), and by \(T\) the set of all stopping times.

I say that an individual is inactive if she is not currently searching for a job, that is, if she has not yet paid the fixed adjustment cost since she started working at her current job. I say that the individual is searching if she has paid the fixed cost but has not yet received an adjustment opportunity. Consider an individual who has just received an adjustment opportunity at time \(0\) when her current wage is \(w_0\), and call \(\hat{V}(w_0)\) her value function at that instant. Her new labor supply \(h_0\) and her next control time \(\tau_1 > 0\) are the solution to the following recursive problem:

\[
\hat{V}(w_0) = \max_{\{h_0, \tau_1\}} \mathbb{E}_0 \left[ \int_0^\infty q e^{-q\hat{\tau}_1} \left\{ \int_0^{\tau_1+\hat{\tau}_1} e^{-(\rho_1+\rho_2)t} U(w_t h_0 - T(w_t h_0), h_0) dt - e^{-(\rho_1+\rho_2)\tau_1} f_{\tau_1} + e^{-(\rho_1+\rho_2)(\tau_1+\hat{\tau}_1)} \hat{V}(w_{\tau_1+\hat{\tau}_1}) \right\} d\hat{\tau}_1 \right],
\]

(7)

subject to the law of motion (2) of the wage. Let \(\mathcal{V}_i(w, h)\), respectively \(\mathcal{V}_s(w, h)\), denote the value function (i.e., the expected present discounted value of lifetime utility net of the adjustment costs) of an inactive, respectively searching, worker with current wage \(w\) and labor supply \(h\). I formally define and characterize the value functions \(\mathcal{V}_i(w, h)\) and \(\mathcal{V}_s(w, h)\) in Section 3.

**Government’s problem.** The government chooses the tax schedule \((\tau, p)\), and evaluates social welfare according to a utilitarian objective over all living individuals. Due to the assumption of exponential deaths, individuals receive equal weights independently of their age. The government maximizes the long-run (steady-state) social welfare subject to a budget balance constraint that imposes that the total tax revenue net of transfer payments must be at least as large as an exogenous revenue requirement \(\bar{R}\). Assuming their existence (see Section 4), let \(f_{w, h}^i(\cdot, \cdot)\) and \(f_{w, h}^s(\cdot, \cdot)\) denote the stationary joint densities of wages \(w\) and hours \(h\) for inactive and searching individuals, respectively. The government solves:

\[
\max_{\{\tau, p\}} \int_0^\infty \int_0^\infty \sum_{x \in \{i, s\}} V_x(w, h) f_{w, h}^x(w, h) dwdh
\]

(8)

subject to

\[
\int_0^\infty \int_0^\infty \sum_{x \in \{i, s\}} T(wh) f_{w, h}^x(w, h) dwdh \geq \bar{R}.
\]

(9)
Let $\lambda$ denote the marginal value of public funds, i.e., the Lagrange multiplier associated with the budget constraint (9). Let $B(T)$ denote the long-run tax revenue given the tax schedule $T(\cdot)$, i.e., the left-hand side of the constraint (9). Finally, let $W(T)$ denote social welfare, equal to the sum of individual indirect utilities (the maximand in (8)), normalized by $\lambda$ to obtain a money metric for welfare.

The goal of this paper is to characterize the optimal tax schedule $T(\cdot) = (\tau, p)$ solution to the government’s problem (8)-(9). I do so in Section 5 by imposing that the first-order effects on social welfare of perturbations $(d\tau, dp)$ of the tax schedule (as $d\tau, dp \to 0$) are equal to zero. The resulting two equations, along with the budget constraint, fully characterize the optimum $(\tau, p, \lambda)$.

3 Individual behavior

In this section I characterize the optimal individual behavior in the model.

Frictionless model. The solution to the frictionless problem (5) is as follows. At each instant $t$, the individual’s optimal labor supply $h^*_t$ is an increasing function of her current wage $w_t$ and her net-of-tax rate $(1 - T'(w_t h^*_t))$. The frictionless taxable income $y^*_t = w_t h^*_t$ and disposable income $c^*_t$ are given by:

\[
y^*_t = (1 - T'(y^*_t))^\varepsilon w_t^{1+\varepsilon} = (1 - \tau)^{\frac{\varepsilon}{1+p\varepsilon}} w_t^{\frac{1+\varepsilon}{1+p\varepsilon}},
\]

\[
c^*_t = y^*_t - T(y^*_t) = \frac{1}{1 - p} (1 - \tau)^{\frac{1+p(1+\varepsilon)}{1+p\varepsilon}}.
\]

Equation (10) shows that the structural elasticity parameter $\varepsilon$ and the parameters of the tax schedule $(\tau, p)$ govern the relationship between an individual’s wage and her corresponding choice of labor supply, taxable and disposable incomes. In particular, the magnitude by which higher wages (or lower marginal tax rates) translate into higher incomes is given by the elasticity $\varepsilon$.

Using equation (10), we find that the laws of motion of the taxable and disposable incomes are given by the following random growth processes:

\[
d\ln y^*_t = \mu_y dt + \sigma_y dW_t, \quad \text{with} \quad \{\mu_y, \sigma_y\} = \frac{1 + \varepsilon}{1 + p\varepsilon} \{\mu_w, \sigma_w\}, \tag{11}
\]

\[
d\ln c^*_t = \mu_c dt + \sigma_c dW_t, \quad \text{with} \quad \{\mu_c, \sigma_c\} = (1 - p) \frac{1 + \varepsilon}{1 + p\varepsilon} \{\mu_w, \sigma_w\}. \tag{12}
\]

Equations (11,12) shows that the income processes are given endogenously from the wage process as functions of the labor supply elasticity $\varepsilon$ and the rate of progressivity $p$ of the tax schedule. In particular, a higher elasticity $\varepsilon$ and a lower progressivity $p$ lead to a higher volatility $\sigma_y$ of the frictionless income process.

---

6More generally, we can use this method to characterize the first-order welfare effects of revenue-neutral tax reforms of any (possibly suboptimal) tax schedule; see Golosov, Tsyvinski and Werquin (2014) for a general theoretical exposition and Section 6 for a numerical analysis along those lines.
I assume that
\[ \rho \equiv \rho_1 + \rho_2 - (1 - \gamma) \mu c - \frac{1}{2} (1 - \gamma)^2 \sigma_c^2 > 0, \]
which will ensure that the individual’s indirect lifetime utility is finite.

**Frictional model.** I now analyze the frictional problem (7) with \( \kappa > 0 \). This problem has two state variables: the wage \( w \) and the labor supply \( h \). The crucial variable for the analysis is the labor supply deviation \( \delta \), which is defined as the log-difference between the actual and the desired (or frictionless) hours of work \( h \) and \( h^* \), that is,
\[ \delta_t \equiv \ln (h_t) - \ln (h_t^*) = \ln (y_t) - \ln (y_t^*), \tag{13} \]
where \( y_t^* \) is given by (10). While the individual remains inactive, her deviation \( \delta_t \) evolves according to the following process:
\[ d\delta_t = -d\ln h_t^* = \mu_\delta dt + \sigma_\delta dW_t, \quad \text{with} \quad \{\mu_\delta, \sigma_\delta\} = -\frac{(1 - p)\varepsilon}{1 + p\varepsilon} \{\mu_w, \sigma_w\}, \tag{14} \]
and upon adjustment of labor supply from \( h \) to \( h' \) at time \( \hat{\tau} \) the deviation jumps from \( \delta_{\hat{\tau}} \) to \( \delta_{\hat{\tau}} + (\ln h' - \ln h) \). In the sequel I change variables and use either \((y^*, \delta)\) or \((y, \delta)\) (rather than \((w, h)\)) as the two state variables of the individual’s problem, i.e., her frictionless or actual taxable income, and the deviation of her income away from its frictionless optimum. There are one-to-one correspondences between these pairs of variables, given by the relationships (10) and (13). Accordingly, with a slight abuse of notation, from now on I denote the individual’s utility function by \( U(y^*, \delta) \) and the value functions by \( V_x(y^*, \delta) \) and \( \bar{V}_x(y, \delta) \equiv V_x(y e^{-\delta}, \delta) \), for \( x \in \{i, s\} \).

We can easily show that the flow utility \( U(y^*, \delta) \) is homogeneous in the utility of frictionless disposable income \( c^* \),
\[ U(y^*, \delta) = \frac{1}{1 - \gamma} \left( \frac{1 - \tau}{1 - p} y^* (1 - p) \right)^{1 - \gamma} \left[ e^{(1 - p)\delta} - \frac{1 - p}{1 + 1/\varepsilon} e^{(1 + 1/\varepsilon)\delta} \right]^{1 - \gamma} \equiv g(c^*) \times u(\delta). \tag{15} \]
A second-order Taylor approximation of the function \( u(\delta) \) around the frictionless optimum \( \delta = 0 \) shows that the utility loss from failing to optimize is locally quadratic,
\[ u(\delta) \approx \left( \frac{1 + p\varepsilon}{1 + \varepsilon} \right)^{1 - \gamma} \left[ 1 - \frac{1}{2} (1 - \gamma) (1 - p) \left( 1 + \frac{1}{\varepsilon} \right) \delta^2 \right]. \tag{16} \]
Since the function \( u(\delta) \) in (15) is not well defined for \( \delta \) far away from 0, I assume for simplicity that the utility of deviation is given by its quadratic approximation (16) for any \( \delta \in \mathbb{R} \). Equation (15) together with the homogeneity of the fixed adjustment cost (6) and the random growth evolution of frictionless disposable income (12) allow us to crucially reduce the dimensionality of the state

\[ \text{Alternatively we can keep the exact expression if } \gamma = 0 \text{ (we should then add curvature to the social welfare function to give the government a redistributive motive). None of the qualitative results would be affected.} \]
space. Specifically, I show the following proposition:

**Proposition 1. (Homogeneity.)** The policy functions and the value functions $V_i(y^*, \delta)$ and $V_s(y^*, \delta)$ of inactive and searching individuals with frictionless taxable income $y^*$ and labor supply deviation $\delta$ are homogeneous of degree one in the utility of desired consumption $g(c^*)$. The value functions can thus be written as

$$V_x(y^*, \delta) = g\left(\frac{1 - \tau}{1 - p} y^*(1-p)\right) v_x(\delta), \quad \forall x \in \{i, s\},$$

for some functions $v_i, v_s : \mathbb{R} \to \mathbb{R}$.

**Proof.** See Appendix. $\square$

I now analyze the solution to the impulse control problem (7). The individual’s optimal adjustment behavior is as follows. For any level of labor supply $h$ (i.e., for any given job), the optimal impulse control policy can be characterized by an interval of inaction $(\hat{\delta}, \bar{\delta})$ and a return point $\delta^*$, with $\delta < \delta^* < \bar{\delta}$. No control is exerted as long as the state process $\hat{\delta}$ is in $(\hat{\delta}, \bar{\delta})$. When the state process strikes or is below $\hat{\delta}$ or above $\bar{\delta}$, the individual pays the fixed cost and waits (on average a duration $q^{-1}$) until she receives an adjustment opportunity. At this time she adjusts the state to $\delta^*$, i.e., hours jump from $h$ to $h' = h \exp(\delta^* - \delta_{\alpha-})$, where $\delta_{\alpha-}$ is the labor supply deviation at the time $\tau$ the signal is received.

I first introduce some useful notation. Define, for any $x \in \{w, y, c, \delta\}$ and $\rho > 0$,

$$r_{1,x}^\rho \equiv \frac{\mu_x}{\sigma_x^2} - \sqrt{\frac{\mu_x^2}{\sigma_x^2} + \frac{2\rho}{\sigma_x^2}}, \quad \text{and} \quad r_{2,x}^\rho \equiv \frac{\mu_x}{\sigma_x^2} + \sqrt{\frac{\mu_x^2}{\sigma_x^2} + \frac{2\rho}{\sigma_x^2}},$$

and for all $\delta \in \mathbb{R}$, letting $\hat{\tau} \sim \mathcal{E}(q)$,

$$\hat{v}_s(\delta) \equiv \mathbb{E}_0 \left[ \int_0^{\hat{\tau}} e^{-(\rho_1 + \rho_2)t} U(y^*_t, \delta_t) dt \right] = \int_{-\infty}^{\infty} \left[ e^{\rho_1 + \rho_2 + \gamma} \mathbb{I}_{x \leq 0} + e^{\rho_1 + \rho_2 + \gamma} \mathbb{I}_{x > 0} \right] u(x + \delta) dx \frac{\sigma_\delta^2}{2} e^{(1-\gamma)(1 + \frac{1}{2})} \left( \frac{\rho_1 + \rho_2 + q}{\rho_2} \delta \bar{\delta} - \frac{\rho_1 + \rho_2 + q}{\rho_1} \right),$$

where the expectation $\mathbb{E}_0$ is given $\delta_0 = \delta$, and the second equality is derived in the Appendix (it follows that $\hat{v}_s(\delta)$ is a quadratic polynomial). For any $n \geq 0$, let $\mathcal{C}^n(\mathbb{R})$ denote the set of $n$-times continuously differentiable functions $v : \mathbb{R} \to \mathbb{R}$. Define the generator $\mathcal{L} : \mathcal{C}^2(\mathbb{R}) \to \mathcal{C}^0(\mathbb{R})$ by

$$\mathcal{L}v(\delta) = [\mu_\delta + (1-\gamma) \sigma_c \sigma_\delta] v'(\delta) + \frac{1}{2} \sigma_\delta^2 v''(\delta),$$

and the intervention operator $\mathcal{M}$ by

$$\mathcal{M}v(\delta) = \frac{q}{\rho + q} \sup_{\delta' \in \mathbb{R}} v(\delta') + \hat{v}_s(\delta) - \kappa.$$

---

See Alvarez and Stokey (1998) for a general theoretical analysis of this result.
The following proposition is a Verification Theorem that provides sufficient conditions for optimality and characterizes the optimal policy:

**Proposition 2. (Verification.)** Suppose we can find a function \( v : \mathbb{R} \to \mathbb{R} \) that satisfies:

(i) (Smooth-fit principle.)

\[
\forall \delta \in \mathbb{R}, \quad v(\delta) \in C^1(\mathbb{R}) \cap C^2(\mathbb{R} \setminus D), \quad \text{where } D = \{ \delta \in \mathbb{R} : v(\delta) > Mv(\delta) \}.
\]  

(ii) (Value-matching QVI.)

\[
Mv(\delta) \leq v(\delta), \quad \forall \delta \in \mathbb{R}.
\]

(iii) (Hamilton-Jacobi-Bellman QVI.)

\[
Lv(\delta) + u(\delta) \leq \rho v(\delta), \quad \forall \delta \in \mathbb{R} \setminus \partial D, \quad \text{with equality in } D.
\]

(iv) (Technical conditions.) \( v \) has locally bounded derivatives near \( \partial D \) and,

\[
E \left[ e^{-\rho_1 - \rho_2 t} |v_\delta^\tau(\tau)| + \int_0^\infty e^{-\rho_1 - \rho_2 t} |Lv_\delta^\tau(\tau)| \, dt \right] < \infty, \quad \forall \tau \in T, p \in P, \delta_0 \in \mathbb{R}.
\]

Then \( v(\delta) = v^i(\delta) \) and \( Mv(\delta) = v^s(\delta) \) for all \( \delta \in \mathbb{R} \), where \( v^i \) and \( v^s \) are the individual’s value functions (63). Moreover, the optimal impulse control policy \( p^* \in P \) is given by (using the notations introduced in Section 3 and assuming that \( \{v(\delta^\tau_\tau) : \tau \in T\} \) is uniformly integrable): for all \( j \geq 1,\)

\[
\tau_j^* = \inf \left\{ t > \tau_{j-1}^* + \tilde{\tau}_{j-1} : \delta_t^{p^*_{j-1}} \notin D \right\}, \quad \text{and } \Delta_j^* = \sup_{\delta \in \mathbb{R}} v(\delta') - v\left(\delta_t^{p^*_{j-1}}(\tau_{j-1}^*) \right),
\]

where \( \delta_t^{p^*_{j-1}} \) is the process resulting from applying \( p_j^* = (\tau_1^*, \ldots, \tau_{j-1}^*, \Delta_1^*(\tau_1), \ldots, \Delta_{j-1}^*(\tau_{j-1})) \) to \( \delta \).

**Proof.** See Appendix.

The optimal policy described above dictates that when the state process \( \delta \) stays within the continuation region \( D \) the control takes no action, and when it attempts to leave the set \( D \) the control instantaneously moves the process to a state within \( D \). Conditions (21) and (22) are the Quasi-Variational Inequalities (QVI) (Bensoussan and Lions (1982)). Note that, as is standard with such Verification results, condition (20) assumes a priori the smooth-fit property through the inaction and search regions.

The method to solve the individual’s optimal control problem then consists of conjecturing the optimal policy and corresponding value function (using the conditions (20,21,22)) and then verifying that the conditions of Proposition 2 are satisfied (including that the QVI hold on the entire domain \( \mathbb{R} \)). Specifically, guess that the optimal impulse control \( p^* \) is defined by the interval of inaction

---

9In general, however, the function \( v \) need not be \( C^1 \) everywhere. To handle this situation we need to work with viscosity solutions to the HJB equation, see, e.g., Chapter 9 in Oksendal and Sulem (2007).
(δ, ̄δ) and the return point ̄δ∗, so that the individual pays the fixed cost at the stopping time τ when the state process δt hits ̄δ or ̄δ, and jumps to ̄δ∗ when she receives an adjustment offer. The value functions vi, vs are then constructed as follows. Consider first an inactive individual, i.e., who has not yet paid the fixed cost since she started her current job. As long as she remains inactive, her value function vi(δ) satisfies the following Hamilton-Jacobi-Bellman (HJB) equation:

\[ \frac{1}{2} \sigma^2 \delta v''_i(\delta) + [\mu_\delta + (1 - \gamma) \sigma_c \sigma_\delta] v'_i(\delta) - \rho v_i(\delta) = -u(\delta), \quad \forall \delta \in (\underline{\delta}, \overline{\delta}), \]  

(24)

subject to the following boundary conditions at ̄δ, ̄δ∗, ̄δ: value-matching, smooth-pasting, and optimality:

\begin{align*}
v_i(\overline{\delta}) &= v_s(\overline{\delta}) - \kappa, \quad \text{and} \quad v_i(\delta) &= v_s(\delta) - \kappa, \\
v'_i(\delta^+) &= v'_i(\delta^-) = v'_s(\delta), \quad \text{and} \quad v'_i(\delta^+) &= v'_i(\delta^-) = v'_s(\delta), \\
v'_i(\delta^*) &= 0,
\end{align*}

(25) (26) (27)

where the value of searching is

\[ v_s(\delta) = \hat{v}_s(\delta) + q \frac{\rho + q}{\rho} v_i(\delta^*). \]

(28)

Note that the differential equation (24) with boundary conditions (25) to (28) defines a free-boundary problem, as the boundaries (\underline{δ}, ̄δ∗, ̄δ) of the inaction region are not fixed but rather must be determined as part of the solution.\(^{10}\) By Proposition 2, the solution \((v_i, v_s, \underline{\delta}, ̄\delta, ̄\delta)\) to this free boundary problem is the value function of the individual problem, and the associated impulse control policy is optimal, if it satisfies the conditions \(\mathcal{L}v + u \leq \rho v\) on \(\mathbb{R} \setminus (\underline{\delta}, \overline{\delta})\) and the \(v_s - \kappa \leq v_i\) on \((\underline{\delta}, \overline{\delta})\), along with condition (iv) of Proposition 2. In the simple case where adjustment is immediate following the payment of the fixed cost (that is, there is no search and the environment reduces to a two-sided \((S, s)\) model), we can show directly that these conditions hold and we obtain:

**Corollary 1.** Let \(q = \infty\), so that \(\hat{v}_s(\delta) = 0\). Suppose that there exist \((\underline{\delta}, \delta^*, \overline{\delta})\) such that the function \(v_i : \mathbb{R} \to \mathbb{R}\) solves the differential equation problem (24), (28), (25), and that the conditions (26) and (27) hold. Suppose that \(v'(\delta) > 0\) (resp., \(v'(\delta) < 0\)) on an interval \(\delta \in (\underline{\delta}, \overline{\delta} + \varepsilon)\) (resp., \(\delta \in (\overline{\delta} - \varepsilon, \overline{\delta})\)). Then \(v_i(\delta)\) is the value function of the individual’s problem and the optimal individual policy is characterized by the control band \(\{\underline{\delta}, \delta^*, \overline{\delta}\}\).

*Proof.* See Appendix.

In the Appendix I also derive heuristically, using dynamic programming arguments, the QVI conditions of Proposition 2. I now describe the intuition for these equations. Consider first an individual who is searching, i.e., who has paid the fixed cost but not yet received the adjustment

\(^{10}\)For a given impulse control policy defined by three (potentially suboptimal) thresholds \(\underline{\delta}, \overline{\delta}^*, \overline{\delta}\), we obtain a fixed boundary problem which can be solved in closed-form by integrating the second-order differential equation (24) and pinning down the solution’s coefficients by inverting the linear system of two equations and two unknowns (25).
offer. By equation (19), the first term in the right hand side of (28) is the flow utility from the time at which the cost is paid until the adjustment occurs, and the second term is the expected value of returning to \((y^*, \delta^*)\), i.e., \(\mathbb{E}\left[ e^{-(\rho_1+\rho_2)\hat{\tau}} \mathcal{V}_i \left( y^*_i, \delta^*_i \right) \right] \). Next consider an inactive individual. The value-matching conditions (25) at the thresholds \(\hat{\delta}, \delta\) mean that at the time the agent decides to pay the fixed cost and search for a new job, she must be indifferent between doing so and continuing with her current job. The smooth-pasting conditions (26) mean that the marginal value and the marginal cost of starting to search must be equal, so that the two value functions \(v_i\) and \(v_s\) meet tangentially at the boundaries of the inaction region. The optimality condition (27) at \(\delta^*\) means that the optimal return point upon adjustment is the maximum of the value function \(v_i(\delta)\) of the newly inactive individual. Finally, the intuition for the HJB equation (24) is as follows. Interpreting the entitlement to the flow of disposable incomes and deviations as an asset, and \(\mathcal{V}_s(y^*, \delta)\) as its value, we can write:

\[
(\rho_1 + \rho_2) \mathcal{V}_i(y^*, \delta) = U(y^*, \delta) + \mathbb{E}_t \left[ d\mathcal{V}_i(y^*, \delta) \right] dt.
\]

The left hand side gives the normal return per unit time that an individual, using \((\rho_1 + \rho_2)\) as the discount rate, would require for holding this asset. The right hand side is the expected total return per unit time from holding the asset. The first term is the immediate payout or dividend from the asset. The second term is its expected rate of capital gain or loss. The equality is a no-arbitrage condition, expressing the investor’s willingness to hold the asset. Using Itô’s formula, we can express the second term in the right hand side as a function of the first and second partial derivatives of the value function \(\mathcal{V}_i\) and the drifts and volatilities of the income and deviation processes. We then obtain the IJB equation (24) for \(v_i(\delta)\) using the homogeneity of the value function shown in Proposition 1.

Figure 2 shows graphically the optimal individual behavior (left panel) and the corresponding value functions (right panel). In the left panel, an individual’s labor supply moves along the horizontal red line as long as she remains inactive; that is, her actual labor supply \(h\) is constant while her desired labor supply \(h^*\) tracks the evolution of her productivity. When she reaches the boundaries of the inaction region (i.e., the thick blue lines \(h^* = e^{-\delta} h\) or \(h^* = e^{-\hat{\delta}} h\)), she starts searching and, as soon as she receives an offer, adjusts up or down to a new labor supply level on the central blue line \(h' = e^{\delta^*} h^*\). The right panel shows the value functions \(v_i(\delta)\) and \(v_s(\delta)\) of inactives (in blue) and searchers (in red) as well as the optimal control band. The value of inactives is bell-shaped and reaches its maximum at \(\delta^*\). When reaching the boundaries \(\hat{\delta}\) and \(\delta\) of the inaction region, the individual becomes a searcher and her value function jumps up along the corresponding dashed blue line. The size of the jump, i.e. the difference between the two functions at these values, is equal to the fixed cost \(\kappa\). Finally, the dashed red curve is the function \(v_s(\delta) - \kappa\), which illustrates that the two value functions meet tangentially at the boundaries of the inaction region (smooth-pasting).
Effects of taxes on individual behavior. I first describe how tax policy affects the frictionless income variables. Equation (10) implies that the effects of perturbing the parameters $\tau, p$ of the tax schedule on the individual frictionless taxable income $y^*$ are given by

$$\frac{d \ln y^*}{d \ln (1 - \tau)} = \frac{\varepsilon}{1 + p\varepsilon}, \quad \text{and} \quad \frac{d \ln y^*}{dp} = -\frac{\varepsilon}{1 + p\varepsilon} \ln y^*. \quad (29)$$

The interpretation of equation (29) is as follows. The behavioral change in income following a tax increase (both in $\tau$ and in $p$) is determined by the structural elasticity parameter $\varepsilon$. If the baseline tax system is linear, i.e. $p = 0$, (29) implies immediately that the elasticity of labor income $y^*$ with respect to the net-of-tax rate $1 - \tau$ is equal to $\varepsilon$. Suppose now that the baseline tax system is non-linear (progressive or regressive), i.e. $p \neq 0$. Then a change in the marginal tax rate $T'(y^*)$ that an individual faces induces a direct reduction of his labor income $y^*$ by the amount $\varepsilon$, by definition of the labor income elasticity. This direct adjustment generates in turn an indirect change in the marginal tax rate that the individual faces, due to the non-linearity of the baseline tax schedule. The amount of this change is equal to $d(T'(y)) = T''(y) dy$, and it induces a further labor income adjustment given by the elasticity $\varepsilon$ and the curvature $p$ of the baseline tax schedule. Thus the total change in income following a perturbation of the net-of-tax rate $(1 - T'(y^*))$ of an individual with income $y^*$ is given by

$$\frac{d \ln y^*}{d \ln (1 - T'(y^*))} = \frac{\varepsilon}{1 + T''(y^*)} \frac{y^*}{1 - T'(y^*)} = \frac{\varepsilon}{1 + p\varepsilon}. \quad (30)$$

Equations (29) and (30) thus show that, from the point of view of individuals, the effect on income of a percent perturbation of the parameter $(1 - \tau)$ is equivalent to a percent perturbation of the net-of-tax rate $(1 - T'(y))$ at every income level. Similarly, the effect of a perturbation of the parameter $p$ is equivalent to perturbing the marginal tax rates faced by all individuals by an amount proportional
to their log-income, so that the magnitude of the tax increase raises with income.

Taxes also affect the laws of motion of the frictionless income variables. Equations (11,12) imply that the effects of perturbing the parameters $\tau, p$ on the drift and volatility of the frictionless taxable and disposable income processes are given by:

\[
\frac{d \ln \{|\mu_y|, |\sigma_y|\}}{d \ln (1 - \tau)} = 0, \quad \text{and} \quad \frac{d \ln \{|\mu_y|, |\sigma_y|\}}{dp} = -\frac{\varepsilon}{1 + p\varepsilon} < 0, \tag{31}
\]

\[
\frac{d \ln \{|\mu_c|, |\sigma_c|\}}{d \ln (1 - \tau)} = 0, \quad \text{and} \quad \frac{d \ln \{|\mu_c|, |\sigma_c|\}}{dp} = -\frac{1}{1 - p} \frac{1 + \varepsilon}{1 + p\varepsilon} < 0.
\]

Thus a higher rate of progressivity of the tax schedule lowers the drift and the volatility of both the taxable and disposable income processes. Intuitively, individual income responses following an increase in productivity are attenuated by the fact that higher incomes pay higher marginal tax rates if the tax schedule is progressive. Note that a uniform change in the marginal tax rates (i.e., a change in $\tau$), on the other hand, does not affect the drift or the volatility of income since all incomes are shifted by a proportional amount.

Next, I turn to the effects of taxes on the optimal individual adjustment policy in the frictional model. The parameter $\tau$ has no effect on the optimal behavior $\{\delta, \delta^*, \delta\}$. Equation (31) shows that a decrease in $p$, however, increases the volatilities $\sigma_y, \sigma_\delta$ of the income and the deviation processes. This in turn raises the option value of waiting to adjust labor supply, and therefore widens the optimal inaction region. Hence

\[
\frac{d \ln \{|\delta|, \delta\}}{d \ln (1 - \tau)} = 0, \quad \text{and} \quad \frac{d \ln \{|\delta|, \delta\}}{dp} < 0. \tag{32}
\]

Intuitively, this is because a less progressive tax schedule magnifies the unexpected shocks to the wage. This raises the incentives for the individual to keep her current job and wait to observe the evolution of her productivity before carrying out the costly adjustment, in order to save on new search costs.\footnote{\textsuperscript{11}This option value effect is similarly at play in models of labor demand where adjustment is characterized by an optimal threshold value, e.g., the Mortensen and Pissarides (1994) model with endogenous job destruction.} Note that a lower rate of progressivity has an ambiguous effect on the frequency of adjustment of labor supply: on the one hand, the higher volatility of the deviation process makes individuals reach the boundaries of their inaction region faster; on the other hand, the inaction region is wider, which tends to make them adjust less often. In practice, however, the volatility effect typically dominates the size-of-the-bands effect, so that a less progressive tax schedule increases the frequency of adjustment. Intuitively, the desired (frictionless) income moves away from the current income faster, so that it is optimal for the individual to carry out the adjustment more often.

**Individual welfare.** In the frictionless model, the value function $V^*(y)$ of an individual with current income $y$ is given by

\[
V^*(y) = \frac{1}{\rho_1 + \rho_2 - (1 - \gamma)} \mu_c - \frac{1}{2} (1 - \gamma)^2 \sigma_c^2 g \left( \frac{p\varepsilon}{1 - p} \frac{1 - \tau}{1 - \gamma} y^{1-p} \right). \tag{33}
\]
Note that in this expression the relevant discount rate \( \rho \) used to compute the present value of utility (the denominator of (33)) depends on the growth rate of the future utility of consumption, and is therefore endogenous to taxes.

In the frictional model, consider an inactive or searching individual with actual (as opposed to frictionless) income \( y \) and deviation \( \delta \). Equation (17) implies that her value function \( \bar{V}_x(y, \delta) \) can be expressed as the value \( \mathcal{V}^*(y) \) that the planner would compute for an individual with income \( y \) wrongly assuming that the world is frictionless, times a correction factor \( \bar{v}_x(\delta) \) which depends on the deviation \( \delta \) of her observed income away from its desired level:

\[
\bar{V}_x(y, \delta) = \mathcal{V}^*(y) \times \bar{v}_x(\delta),
\]

where the value functions \( \bar{v}_i(\delta), \bar{v}_s(\delta) \) are given by

\[
\bar{v}_x(\delta) = \rho \left(1 + \frac{p \varepsilon}{1 + \varepsilon}\right)^{-\gamma - 1} e^{-(1-p)(1-\gamma)\delta} v_x(\delta),
\]

for \( x \in \{i, s\} \). These correction terms \( \bar{v}_x(\delta) \) are strictly decreasing in the deviation \( \delta \), as individuals who earn a given income \( y \) but work fewer hours (i.e., have a higher wage) get a higher utility than those who earn the same income but work more hours at a lower wage. I finally denote by \( \bar{V}(y, \delta) \) and \( \bar{v}(y, \delta) \) the weighted averages of the value functions of inactive and searching individuals, that is,

\[
\bar{v}(y, \delta) = \bar{v}_i(\delta) f_{y,\delta}^i(y, \delta) + \bar{v}_s(\delta) f_{y,\delta}^s(y, \delta), \quad \text{and} \quad \bar{V}(y, \delta) = \mathcal{V}^*(y) \bar{v}(y, \delta),
\]

where \( f_{y,\delta}^i, f_{y,\delta}^s \) denote the stationary joint densities of inactive and searching individuals at income and deviation \( (y, \delta) \) (see Section 4), and \( f_{y,\delta} = f_{y,\delta}^i + f_{y,\delta}^s \).

Unlike the frictionless environment, in which there is a representative agent at each income level \( y \), the labor supply adjustment frictions imply that there is a heterogeneous population of individuals who earn the same income but reach different utility levels – both because their wage-hours bundles vary, and because their employment status (inactive or searching) differ. A key friction in this paper is that the government is restricted to using a tax \( T(\cdot) \) on observed incomes, and hence cannot differentiate between various individuals who earn the same amount but work and search differently. This is why the relevant objects for the government to consider are the averages over hours and employment status of the individual utilities, i.e. (36). Importantly this distribution of utilities within income levels, summarized by the function \( \bar{v} \), depends on the progressivity parameter \( p \) and is thus endogenous to tax policy. I analyze the implications of these observations for optimal taxation in Section 5.

Figure 3 shows graphically the value functions within observed income groups \( \bar{v}_i(\delta), \bar{v}_s(\delta) \) (left panel) and the effects of perturbing the progressivity of the tax schedule on the optimal inaction region and on the value of inactives \( \bar{v}_i \) (right panel). The left panel shows that (both inactive and searching) individuals with higher deviation but the same income are worse off. As in Figure 2, within the inaction region the value of searching is always strictly higher than the value of inactivity.
The right panel shows that as the progressivity decreases (linear versus U.S. tax schedule), the inaction region widens, as discussed above, and the distribution of utilities adjusts endogenously within the new bands.

Figure 3: Value functions conditional on observed income $\bar{v}(\delta)$ and effects of progressivity $p$

4 Aggregation

In this section I characterize the long-run wage and income distributions obtained by aggregating the optimal individual policies described in Section 3.

4.1 Stationary wage distribution

A variable $x$ has a double-Pareto-lognormal distribution (DPLN) (or $\ln x$ has a Normal-Laplace distribution) with parameters $(r_1, r_2, m, s^2)$ if its density is given by

$$f_x(x) = \frac{|r_1|}{|r_1| + r_2} \left\{ e^{\frac{1}{2}r_2^2 x^2 - r_1 m x^2 - 1} \Phi \left( \frac{\ln x - m}{s} + r_1 s \right) + e^{\frac{1}{2}r_2^2 x^2 - r_2 m x^2 - 1} \Phi \left( \frac{\ln x - m}{s} + r_2 s \right) \right\}. \tag{37}$$

The double-Pareto-lognormal distribution closely approximates the actual wage and income distributions observed empirically (see, e.g., Reed (2003), Reed and Jorgensen (2004), Toda (2012)). The following proposition shows that the wage distribution converges to a DPLN stationary distribution:

**Proposition 3.** The distribution of wages $w$ converges towards a unique stationary distribution $f_w(\cdot)$ which is double-Pareto-lognormal with parameters $(r_{1,w}, r_{2,w}, m_w, s_{w}^2)$, where $r_{1,w}, r_{2,w}$ are defined in (18) and $m_w, s_w^2$ are the mean and variance of the wage distribution at birth, $f_{w_0}(\cdot)$. In particular, the stationary wage distribution $f_w(\cdot)$ exhibits power-law behavior in both tails, with
Pareto coefficients on the right and left tail respectively given by \( \left( r_{1,w}^{p_2}, r_{2,w}^{p_2} \right) \), that is,

\[
 f_w(w) \sim w^{r_{w}^{p_2} - 1}, \quad \text{and} \quad f_w(w) \sim w^{r_{w}^{p_2} - 1}.
\]  

(38)


The aggregation of the random growth individual wage processes naturally generates the wage distribution’s Pareto tails (see, e.g., Nirei and Souma (2007), Gabaix (2009)), which is one of the most robust empirical stylized facts (as well as an important determinant of optimal taxes, see Saez (2001)). The wage process (2) therefore fits both the microeconomic empirical evidence (see Section 2) and the macroeconomic properties of the wage distributions. The higher the Pareto coefficient (in absolute value), the thinner the tail, the more equal the distribution. A higher drift \( \mu_w \) and a higher volatility \( \sigma_w \) of individual income, and a lower death (or retirement) rate \( \rho_2 \), lead to a more unequal distribution, i.e., a smaller value of \( \frac{\mu}{\sigma} \).

Note that the frictionless taxable and disposable incomes \( y^*, c^* \) are also log-normally distributed at birth with respective mean and variance \((m_y, s_y)\) and \((m_c, s_c)\) (see Appendix) and follow random growth processes (11,12) from then on. Hence their corresponding stationary distributions \( f_y^*, f_c^* \) are also double-Pareto lognormal with respective parameters \( \left( r_{1,y}^{p_2}, r_{2,y}^{p_2}, m_y, s_y^2 \right) \) and \( \left( r_{1,c}^{p_2}, r_{2,c}^{p_2}, m_c, s_c^2 \right) \).

4.2 Stationary income distributions

I now characterize the stationary joint distributions \( f_{\ln y^*, \delta}^i \) and \( f_{\ln y^*, \delta}^s \) of frictionless taxable incomes \( \ln y^* \) and labor supply deviations \( \delta \) for inactive and searching individuals, respectively. Denote by \( f_1, f_2 \) their partial derivatives with respect to the first and second variables, and by \( f_{11}, f_{12}, f_{22} \) their second partial derivatives. We have \( f^i = 0 \) for all \( \delta < \bar{\delta} \) and \( \delta > \bar{\delta} \). Moreover, for all \( \ln y^* \in \mathbb{R} \), all \( \delta \in (\bar{\delta}, \bar{\delta}^*) \cup (\bar{\delta}^*, \bar{\delta}) \) if \( f = f^i \), and all \( \delta \in \mathbb{R} \setminus \{\bar{\delta}, \bar{\delta}^*\} \) if \( f = f^s \), these distributions are the solutions to the following Kolmogorov-forward (KFE) (or Fokker-Planck) equations:

\[
0 = - (\rho_2 + q I_{f^s}) f - \mu_y f_{11} + \mu \delta f_{12} + \frac{1}{2} \sigma_y^2 f_{111} + \frac{1}{2} \sigma_y^2 f_{22} - \sigma_y \delta f_{12},
\]  

(39)

where \( I_{f^s} \) is equal to one if \( f = f_{\ln y^*, \delta}^s \) and zero if \( f = f_{\ln y^*, \delta}^i \).

The Kolmogorov forward equations (39) have the following interpretation. At a given frictionless income level \( \ln y^* \), the density of individuals with deviation \( \delta \notin \{\bar{\delta}, \bar{\delta}^*, \bar{\delta}\} \) is reduced by those who move away from there, and is increased by those who move to \( \delta \) from a former deviation \( \delta' \neq \delta \), following either an increase in their wage if \( \delta' > \delta \) (so that they would now like to work more), or a decrease in their wage if \( \delta' < \delta \). These flows occur both because of the drift \( \mu_y \) (second and third terms of (39)) and the volatility \( \sigma_y \) (fourth to sixth terms of (39)) of individual incomes and deviations. Moreover, at any point \( (\ln y^*, \delta) \), the distribution loses mass at rate \( \rho_2 \) (due to the movements out of the labor force) plus \( q \) for the individuals who are searching (due to the exogenous adjustment opportunities they receive). In the steady-state, these flows in and out of \( (\ln y^*, \delta) \) must balance on net and are thus equal to zero. Note that these equations do not hold at \( \delta^* \) for \( f_{\ln y^*, \delta}^i \),
and at \( \{ \delta, \bar{\delta} \} \) for \( f_{lny*}^{i}, \) where the inflows from births and from endogenous adjustments produce kinks in the densities.

The boundary conditions of the partial differential equations (39) are the following. First, the density functions \( f_{lny*}^{i} \) and \( f_{lny*}^{s} \) are continuous in \( \delta^{*} \) and \( \{ \delta, \bar{\delta} \} \) respectively: for all \( u \in \mathbb{R}, \)

\[
f_{lny*}^{i,s}(u, \delta) = f_{lny*}^{i,s}(u, \delta^{+}), \quad \text{for} \quad \delta \in \{ \delta, \delta^{*}, \bar{\delta} \} \quad \text{(40)}
\]

Second, the boundaries \( \delta \) and \( \bar{\delta} \) are absorbing for \( f_{lny*}^{i} \), so that there is no mass of inactive individuals at the edges of the inaction region: for all \( u \in \mathbb{R}, \)

\[
f_{lny*}^{i}(u, \delta) = f_{lny*}^{i}(u, \bar{\delta}) = 0. \quad \text{(41)}
\]

Intuitively, this is because individuals who reach a boundary of their inaction region immediately start searching and leave the inaction state. Third, the density of searchers in a given job converges to zero as \( \delta \to \pm \infty \): for all \( u \in \mathbb{R}, \)

\[
\lim_{\delta \to \pm \infty} f_{lny*}^{s}(u, \delta) = 0. \quad \text{(42)}
\]

Fourth, total flows in and out of \( \delta, \delta^{*}, \bar{\delta} \) must balance, which yields three functional equations linking the density functions \( f_{lny*}^{i} \) and \( f_{lny*}^{s} \). Letting \( \hat{f}^{x} \) denote the function \( \frac{\sigma_{x}}{\sigma_{\delta}} f_{1}^{x} + f_{2}^{x} \) for \( x \in \{ i, s \} \), these conditions write: for all \( u \in \mathbb{R}, \)

\[
\hat{f}^{i}(u, \delta^{*}) - \hat{f}^{i}(u, \delta) = \frac{2}{\sigma_{\delta}} (\rho_{2} f_{lny*} (u) + q f_{lny*}^{s} (u)), \quad \text{(43)}
\]

\[
\hat{f}^{i}(u, \bar{\delta}^{*}) + \hat{f}^{s}(u, \bar{\delta}^{*}) - \hat{f}^{s}(u, \bar{\delta}) = 0, \quad \text{(44)}
\]

\[
\hat{f}^{i}(u, \bar{\delta}^{*}) + \hat{f}^{s}(u, \bar{\delta}^{*}) - \hat{f}^{s}(u, \bar{\delta}) = 0. \quad \text{(45)}
\]

These equations equate the inflows and outflows of individuals going from one state (inaction, search, non-participation) into another, following a change in their wage and hence desired hours, the reception of a new job opportunity, or a “birth”. Finally, a normalizing condition imposing that the total mass of individuals in the population is equal to 1 completes the full characterization of the economy’s steady-state:

\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[ f_{lny*}^{i}(u, \delta) + f_{lny*}^{s}(u, \delta) \right] \text{dud}\delta = 1. \quad \text{(46)}
\]

**Proposition 4.** The stationary distributions \( f_{lny*}^{i} \) and \( f_{lny*}^{s} \) of inactive and searching individuals are characterized by equations (40), (41), (42), (43), and (46). If \( q = \infty \) the stationary distributions of taxable and disposable incomes \( f_{y}, f_{c} \) have Pareto right and left tails with respective Pareto coefficients \( \left( r_{1,y}^{\alpha}, r_{2,y}^{\alpha} \right) \) and \( \left( r_{1,c}^{\alpha}, r_{2,c}^{\alpha} \right). \)

**Proof.** See Appendix.
Ignoring the search period for simplicity \( q = \infty \), Proposition 4 shows that the Pareto coefficients of the tails of the taxable and disposable income distributions are given in closed form as a function of the exogenous Pareto coefficients of the wage distribution and the progressivity \( p \) of the tax schedule,\(^{12} \) by

\[
\left\{ r_{1,y}^{\rho_2}, r_{2,y}^{\rho_2} \right\} = \frac{1 + p\varepsilon}{1 + \varepsilon} \left\{ r_{1,w}^{\rho_2}, r_{2,w}^{\rho_2} \right\}, \quad \text{and} \quad \left\{ r_{1,c}^{\rho_2}, r_{2,c}^{\rho_2} \right\} = \frac{1}{1 - p + 1 + \varepsilon} \left\{ r_{1,w}^{\rho_2}, r_{2,w}^{\rho_2} \right\}. \tag{47}
\]

These expressions show that the elasticity of labor income \( \varepsilon \) and the parameter \( p \) of the tax system determine the amount by which inequality in exogenous productivities or wages translates into taxable and disposable income inequality. In particular, the distribution of desired taxable income is more unequal (thicker tail) than the wage distribution, because wage differences are amplified by the positive labor supply elasticity \( \varepsilon \). This is consistent with the findings of Krueger, Perri, Pistaferri and Violante (2009). Importantly, the Pareto coefficients (47) of the income distributions are endogenous to tax policy. In particular, we have \( r_{1,y}^{\rho_2} \leq r_{1,c}^{\rho_2} < r_{1,w}^{\rho_2} \) if \( p \geq 0 \) and \( r_{1,c}^{\rho_2} \leq r_{1,y}^{\rho_2} < r_{1,w}^{\rho_2} \) if \( p \leq 0 \), with strict inequalities if \( p \neq 0 \), so that the distribution of frictionless disposable income is less unequal (thinner tail) than the distribution of desired taxable income if and only if the tax schedule is progressive. Moreover, the Pareto coefficients of the income distributions are increasing in the progressivity \( p \), i.e., both the before-tax and the after-tax income distributions have thinner tails (are less unequal) when the tax schedule is more progressive:

\[
\frac{d \ln \left\{ r_{1,y}^{\rho_2}, r_{2,y}^{\rho_2} \right\}}{dp} = \frac{\varepsilon}{1 + p\varepsilon}, \quad \text{and} \quad \frac{d \ln \left\{ r_{1,c}^{\rho_2}, r_{2,c}^{\rho_2} \right\}}{dp} = \frac{1}{1 - p} \frac{1 + \varepsilon}{1 + p\varepsilon}. \tag{48}
\]

Thus the effect of progressivity on the tails of the after-tax income distribution is stronger than on those of the pre-tax income distribution, i.e., a more progressive tax schedule reduces inequality in after-tax incomes more than it reduces inequality in pre-tax incomes.

Figure 4 summarizes these results graphically and plots the stationary distributions of wages, incomes and labor supply deviations in the economy, as well as the effects of taxes on the income distributions. The top two graphs of Figure 4 show the distribution of taxable incomes (left panel) and disposable incomes (right panel), and how they change when the tax schedule goes from the U.S. rate of progressivity to a linear tax rate. The mean and variance of both distributions are lower when \( p \) is higher; the tails are thinner, to a much larger extent in the case of the disposable income distribution than of the pre-tax income distribution. The bottom left panel of Figure 4 shows the wage, taxable income and disposable income distributions in log-log scale. This representation illustrates clearly the fact that those distributions all have left and right Pareto tails, corresponding to the asymptotic straight lines whose slopes are equal to the Pareto coefficients. The smaller the slopes in absolute value, the more unequal the distribution: the wage (or productivity) distribution is the most equal, the taxable income distribution is the most unequal (due to the positive labor supply elasticity); the inequality of disposable incomes is smaller than that of taxable incomes and

\(^{12}\)The parameter \( \tau \) does not affect inequality at the top because it reduces all incomes proportionally.
closer to that of wages due to the positive rate of progressivity. Finally, the bottom right panel of Figure 4 shows the inactive individuals’ stationary distributions of deviations $\delta$ conditional on income $y$ for several values of $y$, along with the boundaries of the optimal inaction region $(\delta, \delta^*, \bar{\delta})$.

Figure 4: Stationary income distributions

5 Optimal taxation

In this section, I analyze the effects of taxes on long-run social welfare (comparative statics across steady-states) in order to characterize the optimal tax schedule, that is, the solution to the government’s problem (8,9). Before deriving these formulas, I formally define and characterize two sets of key variables for the analysis of tax policy: the labor income elasticities and the marginal social welfare weights.
5.1 Labor income elasticities

I first define the (frictionless) intensive margin labor income elasticity, $\varepsilon^* (y^*)$, as the elasticity of an individual’s frictionless (or desired) taxable income $y^*$ with respect to the net of tax rate. We saw in Section 3 that this elasticity is constant across individuals and is given by the (normalized) structural elasticity parameter $\varepsilon$.

**Definition 1.** The (frictionless) intensive margin labor income elasticity is defined as

$$\varepsilon^* (y^*) \equiv \frac{d \ln y^*}{d \ln (1 - T' (y^*))} = \frac{\varepsilon}{1 + p\varepsilon}, \; \forall y^* \in \mathbb{R}_+, \; (49)$$

where the second equality follows from equations (29), (30).

In a frictionless world ($\kappa = 0$), this variable $\varepsilon^* (y)$ would be equal to the response of the individual’s true income to a change in the marginal tax rates. It could be estimated empirically by observing the magnitude of the increase in income following an increase in statutory net-of-tax rates (see, e.g., Gruber and Saez (2002)). However, when the adjustment of labor supply in response to changes in the marginal tax rates is frictional ($\kappa > 0$), the individual elasticity of actual income (the “micro” elasticity) is in general equal to zero, if the agent has not yet adjusted her income in response to the tax change (short-run elasticity), and is infinite at the time of adjustment since a small tax increase then induces a discrete jump in income. In this environment we can nevertheless define and observe empirically the long-run elasticity of aggregate labor income with respect to a uniform change in net-of-tax rates (the “macro” elasticity). Formally,

**Definition 2.** The macro labor income elasticity $E$ is defined as

$$E \equiv \frac{d \ln \int_{0}^{\infty} y f_y (y) dy}{d \ln (1 - \tau)}, \; (50)$$

where $f_y (\cdot)$ is the stationary density of incomes.

The following proposition provides a neutrality result that characterizes the relationship between the frictionless individual elasticity and the macro elasticity:

**Proposition 5.** The intensive margin elasticity and the macro elasticity of labor income with respect to the marginal tax rates are equal, that is,

$$E = \varepsilon^* (y) = \frac{\varepsilon}{1 + p\varepsilon}, \; \forall y \in \mathbb{R}_+. \; (51)$$

**Proof.** See Appendix. \qed

Proposition 5 shows that in the frictionless model, the long-run aggregate labor income elasticity $E$ is equal to the elasticity of frictionless individual income, even though there is always (even in

\footnote{In this paper, the relevant elasticities are the Hicksian elasticities, since there are no income effects and the analysis focuses on comparative statics across steady-states (permanent tax changes).}
the steady-state) a non-degenerate distribution of individuals with actual incomes $y$ for a given frictionless income $y^\ast$. Intuitively, in the long-run, individuals have had time to fully adjust their behavior to the new tax schedule, and even though in general they do not actually earn their desired income $y^\ast$ at any given moment in time, the individual errors wash out in the aggregate and the magnitude of the long-run aggregate response to the tax change is then driven by the structural elasticity parameter $\varepsilon$. In other words, in the case of a uniform increase in the net-of-tax rates (i.e., a change in the parameter $\tau$), the economy behaves in the long-run as if there were a representative (frictionless) agent at each income level. This result is related to those of Caplin and Spulber (1986) and Rogerson (1985) who argue that frictions at the micro (individual) level are irrelevant at the macro (aggregate) level. Note finally that the result of Proposition 5 allows us to to recover empirically the structural individual elasticity parameter $\varepsilon$ when individual labor supply is lumpy, by estimating the effect of uniform tax changes on long-run aggregate income.

Next, I define three extensive margin labor income elasticity concepts as the effects on the income distribution of changes in the adjustment thresholds $\delta, \delta^\ast, \delta^\parallel$.

**Definition 3.** The extensive margin labor income elasticities $\Xi(y), \Xi^\ast(y), \Xi^\parallel(y)$ are defined as

$$\Xi(y) \equiv \frac{\partial \ln f_y(y)}{\partial \ln \delta}, \quad \Xi^\ast(y) \equiv \frac{\partial \ln f_y(y)}{\partial \ln \delta^\ast}, \quad \Xi^\parallel(y) \equiv \frac{\partial \ln f_y(y)}{\partial \ln \delta^\parallel},$$

(52)

where $f_y(\cdot)$ is the stationary density of incomes.

These elasticities capture the effects on the number of employed workers at income $y$ of percentage variations in the action thresholds. While the standard intensive margin elasticity of Definition 1 affects the purely frictionless part of income, the extensive margin elasticities affect the purely frictional part of income through the endogenous option value of waiting to adjust labor supply. The key difference between these extensive margin elasticities and those typically defined in the literature (e.g., Saez (2002), Jacquet, Lehmann and Van der Linden (2013)) is that in Definition 3 the income thresholds are not exogenously given as in the case of a 0-1 pure participation decision, but instead are endogenously and optimally determined by the individual’s labor supply decisions.

### 5.2 Marginal social welfare weights

The social welfare effects of taxation can be characterized using the notion of marginal social welfare weights (see, e.g., Saez and Stantcheva (2014) for a recent exposition). Intuitively, in a standard frictionless static model, the welfare weight at income $y$ is defined as the increase in social welfare, expressed in terms of public revenue, of distributing an additional unit of consumption uniformly among individuals who earn income $y$. With a utilitarian social objective, this welfare weight is then equal to the individual’s marginal utility of consumption normalized by the shadow value of public funds, $\lambda^{-1}c_0^{\gamma}$. In this section I define formally and generalize the relevant notions of marginal social welfare weights to the dynamic and frictional environment.

---

\(^{14}\)In particular, note that these weights are endogenous.
First consider, in the frictionless model, the effect of giving an additional marginal consumption stream \( \{ \hat{c}_t \}_{t \geq 0} \) to an individual with income \( y \), where \( \hat{c}_t \) evolves stochastically over time according to the same process \( (\mu_c, \sigma_c) \) as the frictionless disposable income \( c^* \). The frictionless marginal social welfare weight at income \( y \), \( \omega^*(y) \), is defined as the change in the present discounted value of her utility (and hence, in utilitarian social welfare) due to this additional consumption stream. It is given by:

\[
\omega^*(y) = \frac{1}{\lambda^*} \left( \mathbb{E}_0 \left[ \int_0^{\infty} e^{-(\rho + \beta)t}u \left( c^*_t + \hat{c}_t - \frac{1}{1 + 1/\varepsilon} \left( h^*_t \right)^{1 + 1/\varepsilon} \right) dt \bigg| y_0 = y \right] \right)_{\hat{c}_0=0}
\]

where \( \lambda^* \) is the marginal value of public funds in the frictionless model and the second equality is proved in the Appendix.

Now, in the frictional model, consider the effect on the welfare of an individual with current income \( y \) and current deviation \( \delta \) of giving her an additional marginal consumption stream \( \{ \hat{c}_t \}_{t \geq 0} \) as described above. These additional units of consumption do not have the same effect on individuals who earn the same income \( y \) but have different deviations \( \delta \) (i.e., different wage-hours bundles) or employment states \( x \in \{ i, s \} \): the change in the individual \( y, \delta, x \)'s utility is the same as if she earned income \( y \) in the frictionless model, times the correction factor \( \bar{v}_x(\delta) \) defined in (35). Since the income tax system treats all individuals with the same income identically, we define the static intensive margin welfare weight \( \omega(y) \) as the long-run social value of distributing the additional stream \( \{ \hat{c}_t \}_{t \geq 0} \) uniformly among all the individuals who earn the same income \( y \), independently of their deviations \( \delta \) or their employment state \( x \). That is,

**Definition 4.** The static intensive margin welfare weight \( \omega(y) \) is defined as

\[
\omega(y) = \lambda^* \omega^*(y) \times \int_{-\infty}^{\infty} \bar{v}(y, \delta) f_{\hat{d}|y}(\delta | y) d\delta,
\]

where \( \omega^*(y) \) is the corresponding frictionless welfare weight defined in equation (53), \( \bar{v}(y, \delta) \) is defined in equation (36), \( f_{\hat{d}|y} = f_{\hat{d}_i|y} + f_{\hat{d}_s|y} \) is the total density of deviations conditional on an actual income \( y \), and \( \lambda \) is the marginal value of public funds in the frictional model.

Second, in the dynamic model a permanent change in progressivity affects not only the levels of current and future consumption, but also the growth rate of utility \( (1 - \gamma) \mu_c + \frac{1}{2} (1 - \gamma)^2 \sigma_c^2 \) through the drift and volatility of the consumption process, and hence the discount rate \( \rho \) that individuals use to compute their present discounted value of utility (see equation (33)). I thus define the dynamic intensive margin welfare weight \( \varpi(y) \) as the effect of a percentage decrease in the discount rate \( \rho \) on the individual’s welfare. In the frictionless model, this is given by

\[
\varpi^*(y) = -\frac{1}{\lambda^*} \frac{\partial \omega^*(y)}{\partial \ln \rho} = \frac{1}{\lambda^*} \frac{1}{1 - \gamma} \left( \frac{1 + p_c}{1 + \varepsilon} \frac{1 - \tau}{1 - p} y^{1 - p} \right)^{1 - \gamma} \rho \frac{1}{1 - \gamma} \left( \frac{1 + p_c}{1 + \varepsilon} \frac{1 - \tau}{1 - p} y^{1 - p} \right)^{1 - \gamma} \frac{1}{2} (1 - \gamma)^2 \sigma_c^2.
\]
Similarly, in the frictional model I define:

**Definition 5.** The dynamic intensive margin welfare weight \( \varpi(y) \) is defined as

\[
\varpi(y) = \frac{\lambda^*}{\lambda} \varpi^*(y) \times \int_{-\infty}^{\infty} \bar{v}(y, \delta) f_{\delta|y}(\delta|y) \, d\delta,
\]

where \( \varpi^*(y) \) is the corresponding frictionless welfare weight defined in equation (55).

Third, paralleling the discussion leading to Definition 3, I define the extensive margin welfare weights \( \Omega(y), \Omega^*(y), \bar{\Omega}(y) \) as the effects of changes in the thresholds \( \tilde{\delta}, \delta^*, \tilde{\delta} \) on the social welfare at the income level \( y \):

**Definition 6.** The extensive margin welfare weights \( \{\Omega_i(y)\}_{1 \leq i \leq 3} = \{\Omega(y), \Omega^*(y), \bar{\Omega}(y)\} \) are defined as

\[
\Omega_i(y) = \frac{1}{\lambda} \int_{-\infty}^{\infty} \frac{\partial \ln f_{y,\delta}(y, \delta)}{\partial \ln |\delta_i|} \bar{V}(y, \delta) f_{\delta|y}(\delta|y) \, d\delta, \quad \forall i \in \{1, 2, 3\},
\]

where \( \{\delta_i\}_{1 \leq i \leq 3} = \{\tilde{\delta}, \delta^*, \tilde{\delta}\} \) and \( \bar{V}(y, \delta) \) is defined in equation (36).

Finally, we saw in Section 3 that a change in progressivity affects the composition of each income group, and hence social welfare, through the endogenous value function \( \bar{v}(y, \cdot) \). I thus define the composition margin welfare weight \( \hat{\Omega}(y) \) as:

**Definition 7.** The composition margin welfare weight \( \hat{\Omega}(y) \) is defined as

\[
\hat{\Omega}(y) = \frac{1}{\lambda} \int_{-\infty}^{\infty} \left( \frac{\partial \ln \bar{v}(y, \delta)}{\partial p} + \frac{\partial \ln \bar{v}(y, \delta)}{\partial \ln y} \frac{\partial \ln y}{\partial p} \right) \bar{V}(y, \delta) f_{\delta|y}(\delta|y) \, d\delta,
\]

where \( \frac{\partial \ln y}{\partial p} \) is given by (29).

The welfare weight \( \hat{\Omega}(y) \) captures the endogenous change in the equilibrium distribution of utilities \( \bar{v}(y, \delta) \) within the income group \( y \) in response to a change in progressivity. In particular, when \( q = \infty \), i.e. as the model reduces to the two-sided \((S, s)\) environment, \( \hat{\Omega} \) is simply equal to

\[
\hat{\Omega}(y) = \frac{1}{\lambda} \partial \mathbb{E} \left[ \bar{v}(\delta) | y \right] \frac{\partial \mathbb{E}}{\partial p} \nu^*(y),
\]

that is, the change in the average welfare at income \( y \) due to a change in progressivity.

### 5.3 Optimal tax schedule

I now characterize analytically the optimal tax schedule in terms of the labor income elasticities and the social marginal welfare weights defined in Sections 5.1 and 5.2.

The first equation characterizing the optimal tax schedule is obtained by imposing that a perturbation of the parameter \( \tau \) by \( d\tau \) has no first-order effect on social welfare. This equation can
be thought of as pinning down the marginal value of public funds $\lambda$ given a tax schedule $(\tau, p)$.

Intuitively, $\lambda$ (the Lagrange multiplier associated with the constraint (9)) is equal to the social value of redistributing a dollar of tax revenue through a decrease in $\tau$ by $d\tau$, i.e., through a uniform increase in the net of tax rates (taking into account the behavioral responses that this distribution induces).

**Proposition 6.** The optimal tax schedule $(\tau, p)$ satisfies

$$0 = 1 - \int_0^\infty T'(y) \frac{y\partial_\tau(y)}{1 - T'(y)} \varepsilon^*(y) \frac{f_y(y)}{\mathbb{E}T_\tau} dy - \int_0^\infty \partial_\tau(y) \omega(y) \frac{f_y(y)}{\mathbb{E}T_\tau} dy,$$

where $\partial_\tau(y) = \frac{1}{1 - p} y^{1-p}$ and $f_y(y)$ is the stationary density of incomes given the tax schedule $(\tau, p)$. In the frictionless model, the same equation characterizes the optimal tax schedule, except that the marginal social welfare weights $\omega(y)$ are replaced by their frictionless counterparts $\omega^*(y)$.

**Proof.** See Appendix.

The interpretation of equation (59) is as follows. It imposes that at the optimum tax schedule, a perturbation $d\tau$ of the marginal tax rates should have no first-order effects on social welfare. The first and second terms on the right hand side measure the actual change in government tax revenue of a one-dollar statutory increase in taxes, taking into account the induced change in individual behavior. The additional tax liability levied at the income level $y$ after the tax reform is implemented is given, to a first order in $d\tau \to 0$, by $\partial_\tau(y) d\tau$, and the marginal tax rate changes by $\partial_\tau(y) d\tau$. The first term in the right hand side of (59) is the *mechanical effect* of the perturbation, i.e., the statutory increase in government revenue absent behavioral responses. It is equal to $\mathbb{E}[\partial_\tau(y) d\tau] = d\tau \int_0^\infty \partial_\tau(y) f_y(y) dy$, and I normalize the magnitude of the perturbation so that this mechanical effect is equal to one (dollar). The second term in the right hand side of equation (59) is the *behavioral effect* of the perturbation. The increase $dT' \equiv \partial_\tau(y) d\tau$ in the marginal tax rate of an individual with income $y$ induces her to decrease her taxable income by $y \frac{\varepsilon^*(y)}{1 - T'(y)} T'(y)$. This behavioral income response generates a loss in government revenue proportional to the marginal tax rate $T'(y)$. Summing over individuals using the density of incomes $f_y(y)$ yields the second term in (59). Finally, the third term in (59) is the *welfare effect* of the perturbation, expressed in monetary units. An increase in the tax liability of individual $y$ by $dT \equiv \partial_\tau(y) d\tau$ directly reduces her utility and hence social welfare by $\omega(y) \times dT$, by construction of the marginal social welfare weights (53).

This equation (59) has a structure that is identical to those of the “sufficient statistic” formulas derived by, e.g., Saez (2001) and Golosov, Tsyvinski and Werquin (2014) to characterize the optimal tax systems in frictionless models. Note in particular that all the variables other than the marginal

---

15This will be useful in the numerical exercises of Section 6 when I compute the social welfare effects of perturbing the U.S. tax code (rather than the full optimum tax schedule). In this case the marginal value of public funds entering the main formula (61) through the welfare weights will be defined by (59).

16Saez (2001) derived a similar formula for a specific perturbation that is different from $\partial_\tau$. Golosov, Tsyvinski and Werquin (2014) extend his results to general perturbations and obtain formulas of the same form as (59).
social welfare weights in equation (59) (elasticities, tax schedule, income distribution) are empirically observable.\footnote{However these variables are endogenous to the tax system and should be evaluated at the optimum. The values estimated given the current tax code (in particular, the current U.S. income distribution) can be nevertheless used to quantify the welfare effects of local tax reforms, given by the right hand sides of equations (59) and (60,61). See Golosov, Tsyvinski and Werquin (2014) and Section 6 below.}

Proposition 6 extends the result of Proposition 5. Its main insight is that the long-run effects on social welfare of a uniform change in marginal tax rates are the same as we would calculate by assuming that the economy is frictionless and thus has a representative agent at each income level $y$: in particular, as in Proposition 5, the relevant elasticity is the individual frictionless income elasticity $\varepsilon^*(y)$ defined in (49), even though individual labor supply is lumpy and the micro elasticity is never actually equal to $\varepsilon^*(y)$: individual-level frictions wash out in the long-run of the aggregate economy. There is one difference, however, between the frictionless and the frictional versions of the optimal tax formula (59): the frictional marginal social welfare weights $\omega^*(y)$ must be computed by taking into account the non-degenerate distribution of utilities $E[\bar{v}(y, \delta) | y]$ within income groups. In general this correction term varies with income $y$, so that the schedule of frictional welfare weights is not perfectly homothetic to the schedule of frictionless weights, and the effective redistributive tastes of the government have to be adjusted relative to a model with a representative agent at each income level.

Next, I derive the second equation characterizing the optimal tax schedule by imposing that a perturbation of the rate of progressivity $p$ by $dp \to 0$ has no first-order effects on social welfare. The next proposition is the main theoretical result of the paper:

**Proposition 7.** In the frictionless model, the optimal tax schedule $(\tau, p)$ is fully characterized by (9), (59), and

$$0 = 1 - \int_0^\infty T'(y) \frac{y \partial_p \rho(y)}{1 - T'(y)} \varepsilon^*(y) \frac{dF_y(y)}{E_\partial_p} - \int_0^\infty \left[ \partial_p(\omega^*(y)) + \frac{d\ln \rho}{dp} \right] \frac{dF_y(y)}{E_\partial_p},$$

where $\partial_p(\omega) = \left( \ln y - \frac{1}{1-\tau} \right) \frac{\varepsilon}{1-\tau} y^{1-p}$ and $f_y(\cdot)$ is the stationary density of incomes given the tax schedule $(\tau, p)$. In the frictional model, the optimal long-run tax schedule is given by (9), (59), and

$$0 = 1 - \int_0^\infty T'(y) \frac{y \partial_p \rho(y)}{1 - T'(y)} \varepsilon^*(y) \frac{dF_y(y)}{E_\partial_p} - \int_0^\infty \left[ \partial_p(\omega^*(y)) + \frac{d\ln \rho}{dp} \right] \frac{dF_y(y)}{E_\partial_p} + \int_0^\infty \sum_{i=1}^3 \frac{d\ln |\delta_i|}{dp} \Omega_i(y) + \hat{\Omega}(y) \frac{dF_y(y)}{E_\partial_p},$$

where $I$ denote $\{\delta_i\}_{1 \leq i \leq 3} = \{\hat{\delta}, \delta^*, \delta\}$, $\{\Xi_i\}_{1 \leq i \leq 3} = \{\Xi(y), \Xi^*(y), \Xi(y)\}$, and $\{\Omega_i\}_{1 \leq i \leq 3} = \{\Omega(y), \Omega^*(y), \hat{\Omega}(y)\}$, and where $A(y) = T'(y) y E[\delta | y] - \frac{1}{x} E \left[ \frac{\partial \bar{v}(y, \delta)}{\partial \ln y} \delta | y \right]$.\

**Proof.** See Appendix. \qed

The first result of Proposition 7, equation (60), characterizes the optimal tax schedule in the
frictionless model. Its interpretation is similar to that of equation (59), with one important difference. It equates the first-order welfare effects of a perturbation \( dp \) of the optimal tax schedule \((\tau, p)\) to zero. The tax reform induces a change in the tax liability levied at the income level \( y \) given, to a first order in \( dp \to 0 \), by \( \partial_p (y) \, dp \), and the marginal tax rate changes by \( \partial'_\tau dp \). The first term in the right hand side of (60) is the mechanical effect of the tax change, the second term is the behavioral effect (i.e., the loss in government tax revenue for a one-dollar statutory increase in taxes through an increase in progressivity), and the third term is the welfare effect (expressed in monetary units). Importantly, the welfare effect of an increase in progressivity depends on a new welfare weight, \( \varpi^* (y) \), which is absent from expression (59) and of the optimal tax formulas typically derived in static models (Diamond (1998), Saez (2001)). This is because progressivity affects the growth rate of consumption and hence the discount rate \( \rho \) used to compute the present value of utility (hence the term \( d\ln \rho \) multiplying \( \varpi^* (y) \)). This implies that the standard static Mirrlees model, often loosely argued to characterize long-run optimal taxes, fails to capture the true welfare effects of permanent changes in progressivity if the “long-run” is not properly modeled as the steady-state of a dynamic economy.

The second and main result of Proposition 7 is the derivation of equation (61) which characterizes the optimal long-run tax schedule in the frictional model. The first line of this expression (with the exception of the unimportant term \( A (y) )^{18} \) is exactly the same as the frictionless formula (60) (replacing the frictionless welfare weights with their frictional counterparts \( \varpi (y), \varpi (y) \)). The presence of the second line in formula (61) implies that at least in theory, the frictionless formula does not correctly account for all of the long-run effects of taxes when individual labor supply is subject to frictions. There are several new long-run effects of raising progressivity, both on welfare and on revenue.

The differences between the frictionless and frictional formulas (60) and (61) are the presence in the latter of new behavioral (revenue) effects containing the extensive margin labor income elasticities \( \Xi (y), \Xi^* (y), \Xi (y) \), new welfare effects containing the extensive margin welfare weights \( \Omega (y), \Omega^* (y), \Omega (y) \), and a new welfare effect containing the composition margin welfare weight \( \Omega (y) \).

We saw in Definitions 3 and 6 that a change in the optimal individual impulse control policy, i.e., in the option value of waiting to adjust labor supply, affects the density of incomes and the welfare at each income level \( y \). Formally, equation (61) shows (through the terms \( \frac{\partial \ln |\delta_i|}{dp} - \frac{\partial \ln |\sigma_j|}{dp} \)) that these extensive margin effects on revenue and welfare cannot be ignored theoretically unless an increase

---

18 This term appears because because a change in progressivity affects the relative processes driving the frictionless income variables \( \{\mu_y, \sigma_y, m_y, s_y\} \) (with elasticity \( \frac{d\ln \sigma_y}{dp} \)) and the deviation (or hour) variables \( \{\mu_\delta, \sigma_\delta, \hat{\delta}, \delta^*, \hat{\delta} \} \) (with elasticity \( \frac{d\ln \sigma_\delta}{dp} \)), and this decoupling affects the long-run density of incomes; see Appendix for more details. Note that if the fixed adjustment cost were on income \( y \) rather than hours \( h \) (see footnote 4), the volatility of deviations \( \sigma_\delta \) would always be equal to (minus) that of frictionless incomes \( \sigma_y \), and this term would disappear from the optimal tax formula (61); see the earlier version of this paper (Chapter 1 in Werquin (2015)). Moreover, this term has negligible magnitude in practice; see Section 6.

19 Note that while the intensive margin elasticities are multiplied by the marginal tax rate \( T' (y) \) to obtain the behavioral (revenue) effect of the tax change, because an infinitesimal change in income \( dy \) reduces tax revenue by \( T'' (y) dy \), the three extensive margin terms (corresponding to the three variables \( \hat{\delta}, \delta^*, \hat{\delta} \) of discrete adjustment), on the other hand, are multiplied by the average tax rate \( T (y) \).
in the progressivity $p$ of the tax code induces an equal reduction of the volatility of deviations and of the size of the individual inaction region, i.e.,

$$\frac{d \ln \{|\delta|, |\delta^*|, \bar{\delta}\}}{dp} = \frac{d \ln |\sigma_{\delta}|}{dp} = -\frac{1}{1 - p} \frac{1 + \epsilon}{1 + p\epsilon} dp. \tag{62}$$

This condition is satisfied in the case of a uniform change in marginal tax rates, because $\tau$ affects neither the volatility nor the optimal inaction region; this explains why the extensive margin terms do not appear in the formulas of Propositions 5 and 6. In general, however, condition (62) is violated: as we already saw in Section 3, an increase in progressivity both reduces the volatility of idiosyncratic shocks and narrows the inaction region, but the two effects do not exactly cancel out: the volatility effect typically dominates the size-of-the-bands effect. It follows that as soon as the option value of adjusting labor supply is endogenous, the extensive margin terms are non-zero and the effects of taxes on government revenue are not accurately described by the intensive margin terms captured by the variables $\varepsilon^*(y)$ and $\omega(y)$.

Note moreover that even when condition (62) holds, so that the extensive margin revenue and welfare effects are equal to zero, the fact that labor supply is frictional implies that progressivity has non-zero composition margin welfare effects, captured by the welfare weights $\tilde{\Omega}(y)$. From equation (58) (to simplify the discussion consider the case $q = \infty$), we obtain that these weights are in general non-zero, unless in addition to (62) we have $\frac{\partial E[\tilde{v}(\delta)|y]}{dp} = 0$, i.e., if the average welfare within the population earning income $y$ is exogenous to taxes. But this condition is generally not satisfied (see the right panel of Figure 3), because $p$ affects the curvature of the quadratic loss function $u(\delta)$ in equation (16). Thus progressivity has an additional effect on social welfare by affecting the steady-state distribution of heterogeneous utilities within income groups.

I provide further results about the qualitative direction and the quantitative magnitude of these novel effects in the numerical exercises in Section 6. We can already anticipate intuitively the signs of (some of) these effects. The extensive margin effects on welfare $\Omega, \Omega^*, \tilde{\Omega}$ will tend to reduce the gains of raising progressivity (and hence will imply a less progressive optimal tax schedule). This is because we saw that the standard elasticities and welfare weights capture the true welfare effects of taxes if the narrowing of the inaction region (option value effect) exactly compensates the decrease in the volatility of incomes due to an increase in progressivity. In general, however, the former effect is dominated by the latter, so that an increase in progressivity is equivalent to a wider dispersion of individual incomes around their desired values, relative to the frictionless benchmark. This in turn adversely affects welfare.

Extensive margin effects have previously been introduced in models optimal taxation by Saez (2002) and Jacquet, Lehmann and Van der Linden (2013), who derive optimal tax formulas in static frictionless models with a 0-1 decision whether to participate in the labor force. Here, I generalize 20 In the Calvo limit as $\kappa \to 0$ with $q > 0$, the frequency of adjustment is exogenous to taxes by construction, yet (62) is not necessarily satisfied because the optimal return threshold $\delta^*$ may not satisfy the condition. In the alternative version of the model where the fixed cost is on labor income, I showed in Werquin (2014) that we then have $\frac{d \ln |\delta^*|}{dp} = \frac{d \ln |\sigma_{\delta}|}{dp}$, implying that the extensive margin terms are equal to zero when labor supply adjustments are exogenous in the sense of Calvo.
their insights by deriving a formula where the extensive margins of adjustment occur conditional on participation and the thresholds are optimally and endogenously chosen given the fixed cost of adjusting labor supply. Proposition 7 shows that these extensive margin effects matter theoretically even in the long-run, even though by then all individuals will have had the time to adjust their labor supply to the new tax schedule. This is because taxes affect the option value of adjusting the actual labor supply, and not only the optimal desired labor supply. This is a stronger result than those previously derived in the literature that analyze optimal policy in models with an extensive margin or fixed costs (e.g., Saez (2002), Chetty, Looney and Kroft (2009)). Indeed, these models are typically static, so that in response to the tax change a fraction of the population finds it optimal to adjust their labor supply, while the rest of the population never does. This is occurring in my dynamic model only in the short-run: eventually individuals will all optimally adjust their behavior to the new tax code as their characteristics evolve over time. However, Proposition 7 shows that extensive margin effects matter even then, as the adjustment behavior is endogenously determined. These novel forces are not captured by the standard Mirrleesian static taxation frameworks with labor supply responses on the intensive or the extensive margins.

This endogenous extensive margin conditional on participation is important empirically, as the literature repeatedly found that the labor supply adjustments in response to productivity, wage or tax changes are frictional as modeled in this paper, whether the fixed cost represents the search cost of switching jobs (hours constraints within a firm: Altonji and Paxson (1992), Chetty, Friedman, Olsen and Pistaferri (2011)) or cognitive costs (Chetty, Looney and Kroft (2009), and Gelber, Jones and Sacks (2013)). It would be straightforward to add explicitly a participation margin into the model (see for example Alvarez, Borovickova and Shimer (2015)). However, empirically, much of the difference in labor supply across countries with different tax regimes is driven by hours worked conditional on employment (see Davis and Henrekson (2005), and Chetty, Guren, Manoli and Weber (2011)), so that this extension is not crucial for the purposes of this paper.

Moreover, this option value effect would appear in a similar fashion in labor demand models à la Mortensen and Pissarides (1994) with endogenous job destruction shocks: the job destruction cutoff would similarly respond endogenously to the volatility of idiosyncratic shocks, and hence to the progressivity of taxes. (See Davis, Faberman, Haltiwanger, Jarmin and Miranda (2010) for an empirical evaluation of this mechanism.) Thus, the extensive margin implications highlighted in this paper will arise more generally as soon as adjustments of hours of work to shocks are endogenous, whether it is through individual labor supply decisions, or through firm labor demand (firing) decisions. Here, I derive the implications of these effects in a model where there is in addition an intensive margin dimension of labor supply (work intensity or hours), allowing to use the standard taxation framework as a benchmark.

Finally, the results I derive in this paper can be applied more broadly than to this taxation environment. The key finding of this paper is that non-linear policies interact with fixed adjustment costs (i.e., lumpy optimal behavior) to yield real effects in the long-run. This is in contrast with a large literature with fixed costs at the micro level, where the aggregate economy often behaves in the long-run as a frictionless (representative agent) model. For instance, monetary policy is generally
neutral in the long-run, as well as in the short-run in specific settings as in the aggregation of Caplin and Spulber (1987). Similarly, in the public finance context, Rogerson (1988) shows in a model of lotteries that the aggregate economy can behave as a representative agent with a large labor supply elasticity even though the individual (micro) elasticity is equal to zero, so that individual frictions are irrelevant in the aggregate or in the long-run. Chetty, Friedman, Olsen and Pistaferri (2011) reach a similar conclusion in a model with fixed costs. I show in this paper that this insight is correct only if we consider linear policies, which do not affect the idiosyncratic volatility and thus the option value of adjusting behavior (Propositions 5 and 6). The curvature of policy matters in the long-run aggregate (for both efficiency and welfare) when adjustment at the individual level is frictional. This insight should apply generally to models where non-linear policy instruments are available in environments with fixed costs and lumpy behavior.

6 Quantitative analysis

In this section I calibrate the model analyzed in the previous sections, and evaluate quantitatively the effects highlighted in Proposition 7. I first compute the magnitudes of the labor income elasticities and the marginal social welfare weights, given the current U.S. tax code, in a calibrated version of the model. I finally compute the welfare effects of raising the progressivity of the U.S. tax schedule and the qualitative and quantitative errors made in the evaluation of these effects by wrongly assuming that the economy is frictionless.

6.1 Calibration

I calibrate the marginal tax rates and the rate of progressivity \((\tau, p)\) of the tax schedule in the U.S. using the empirical estimates from PSID data of Heathcote, Storesletten, and Violante (2014): \(\tau = -3\) and \(p = 0.151\). This value of \(p\) implies that earning twice as high an income leads to a 15.1\% decrease in the net-of-tax rate (see equation (4)). These parameters yield a value for total U.S. government revenue \(\bar{R} = $2.33tn\)\(^{21}\) which I keep constant throughout the numerical analysis, so that whenever I vary the progressivity \(p\) I adjust \(\tau\) in order to keep the revenue \(\bar{R}\) unchanged (equation (9)).

The theoretical analysis above requires the coefficient of risk aversion \(\gamma\) to be strictly below 1; I take \(\gamma = 0.9\). There is substantial controversy in the literature about the value of the taxable income elasticity \(\varepsilon\). The micro literature typically finds values lower than 0.3, while the macro literature and some structural estimates find it to be closer to 1 (see Saez, Slemrod, and Giertz (2012), and Keane and Rogerson (2012), for an overview of the two strands). In a frictionless environment Gruber and Saez (2002) find an elasticity between 0.4 and 0.6, while in a context closely related to this paper’s model, Chetty (2012) estimates the structural parameter (Hicksian intensive margin elasticity) \(\varepsilon = 0.33\) using a meta analysis of micro and macro studies and allowing for adjustment frictions to reconcile the wide range of estimates. In my baseline calibration I thus take \(\varepsilon = 0.33\). I

---

\(^{21}\)This is computed for a population of 320m.
discuss below the effects of varying the structural elasticity parameter \( \varepsilon \) on the results, in particular, I show how the results are affected for \( \varepsilon = 1 \).

I calibrate the Pareto coefficients of the (observed) U.S. income distribution, \( r_{y,1}^{\rho_2} \) and \( r_{y,2}^{\rho_2} \), given the parameters of the U.S. tax schedule. The Pareto coefficient of the right tail, \( r_{y,1}^{\rho_2} \), is well known: it varies around 2 and has been decreasing (the tail of the distribution has become thicker, i.e. more unequal) in the past few decades. The coefficient of the left tail has been estimated by, e.g., Reed (2003), Reed and Jorgensen (2004), Toda (2012). I take \( \left(r_{y,1}^{\rho_2}, r_{y,2}^{\rho_2}\right) = (-1.9, 1.4) \).

The mean \( m_y \) and variance \( s_y^2 \) of the lognormal "bulk" of the frictionless income distribution are calibrated using the mean and variance of the observed U.S. distribution of log-incomes.\(^{22}\) Using \( \mathbb{E} [\ln y] = 10.3 \) and \( \mathbb{V} [\ln y] = 1 \) I obtain \( (m_y, s_y) = (10.46, 0.43) \).

There is a large literature estimating log-income dynamics that follow a geometric random walk, that is equation (2) without the jumps, see e.g. Meghir and Pistaferri (2004, 2011). The volatility of idiosyncratic wage risk \( \sigma_y^2 \) in my model corresponds to the variance of the permanent component of the individual log-income process in this literature. I calibrate \( \sigma_y^2 = 0.01 \) (see also Jones and Kim (2014), for an estimate in a frictionless model similar to this paper and further references to the empirical literature).\(^{23}\)

Next, note that the cross-sectional income distribution in the economy, specifically the values of the two Pareto coefficients at the tails, allows us to infer information about the time-series of individual income, since these Pareto tails are generated by the underlying random growth process for income. We have

\[
\rho_2 \rho_{y,1} = \frac{2 \mu_y}{\sigma_y^2}, \quad \text{and} \quad \rho_{y,1} \rho_{y,2} = -\frac{2 \rho_2}{\sigma_y^2},
\]

which pin down the drift \( \mu_y \) and the death rate \( \rho_2 \). Note that this leads to a negative drift of income \( \mu_y \), but the growth rate \( \mu_g + \frac{1}{2} \sigma_g^2 \) is positive. I take a discount rate \( \rho_1 \) so that \( (1 + \rho_1 + \rho_2)^{-1} = 0.95 \).

The parameters of the individual wage and consumption processes, \( (\mu_w, \sigma_w) \) and \( (\mu_c, \sigma_c) \), and those of the wage and consumption distributions, \( \left(m_w, s_w, \rho_{1,w}^{\rho_2}, \rho_{2,w}^{\rho_2}\right) \) and \( \left(m_c, s_c, \rho_{1,c}^{\rho_2}, \rho_{2,c}^{\rho_2}\right) \) are then obtained from equations (12) and (47). In the numerical exercises below, I compute the effects of taxes keeping the parameters of the exogenous wage (or productivity) process \( (\mu_w, \sigma_w, m_w, s_w) \) constant, and use equations (11,12,47) to infer those of the endogenous income distribution.

The fixed adjustment cost \( \kappa \) and the arrival rate of costless adjustment opportunities \( q \) are calibrated as follows. I take the average duration of searching for a new job, equal to \( t_s = 1 \) month, and the average duration of a job (with a given amount of hours) equal to \( t_i + t_s = 5 \) years. Using the explicit expressions

\[
t_s = q^{-1}, \quad \text{and} \quad t_i = \frac{\overline{\delta} - \delta}{\mu_y} \left[ \frac{\delta^* - \delta}{\overline{\delta} - \delta} - e^{2 \bar{\delta} \mu_y / \sigma_y^2} - e^{2 \delta \mu_y / \sigma_y^2} \right],
\]

\(^{22}\)In the frictionless model, these coefficients are given in closed form by \( \mathbb{E} [\ln y] = m_y - r_{y,1}^{-1} - r_{y,2}^{-1} \) and \( \mathbb{V} [\ln y] = s_y^2 + r_{y,1}^{-2} + r_{y,2}^{-2} \).

\(^{23}\)If the analysis is extended to allow for jumps in the wage process, the corresponding parameters can be calibrated from Guvenen, Karahan, Ozkan and Song (2014) who find a double-Pareto distribution of earnings growth rates \( f_v \).
I obtain the values of \( \kappa \) and \( q \) that yield these average durations.

For \( \varepsilon = 0.33 \), I get \( \kappa = 0.0038 \). This value implies that the cost of searching for a new job, \( \kappa \), is equal to 0.38\% of the instantaneous utility \( g(c_0^\ast) \), or 1.2\% of the average monthly utility

\[
\mathbb{E} \int_0^T e^{-(\rho_1+\rho_2)t} g(c_t^\ast) \, dt = \rho^{-1} \left( 1 - e^{-\rho T} \right) \left( \frac{1+\rho_2}{1+\varepsilon} \right)^{1-\gamma} g(c_0^\ast),
\]

that is, 1.2\% of the total utility received during the duration of the (one-month long) search. Despite this relatively small value for the fixed cost, the corresponding inaction region is large (because the utility loss from choosing hours suboptimally is second-order, see equation (16)) and given by \( \delta = -0.09 \) and \( \bar{\delta} = 0.09 \), so that an individual starts searching for a new job when her hours are approximately 9\% away from their optimal value (given her wage). Finally, she then adjusts to \( \delta^\ast = -0.001 \), i.e. 0.1\% below her current optimal value (because of the small drift \( \mu_\delta \)). For \( \varepsilon = 1 \), I get \( \kappa = 0.015 \), which implies the cost of searching for a new job \( \kappa \) is equal to 5.1\% of the average monthly utility. The corresponding inaction region is given by \( \delta = -0.19 \) and \( \bar{\delta} = 0.19 \), so that an individual starts searching for a new job when her hours are about 20\% away from their optimal value.

6.2 Numerical results

I first compute the extensive margin labor income elasticities around the current U.S. tax code, using Definition 3. They are represented in Figure 5 for \( \varepsilon = 0.33 \) on the left panel and \( \varepsilon = 1 \) on the right panel). Figure 5 plots the extensive margin elasticities \( \Xi, \bar{\Xi} \) weighted by the effect of progressivity on the adjustment thresholds, i.e., the tax elasticities

\[
\left( \frac{d \ln |\delta|}{dp} - \frac{d \ln |\sigma_\delta|}{dp} \right) \Xi(y), \quad \left( \frac{d \ln \delta}{dp} - \frac{d \ln |\sigma_\delta|}{dp} \right) \bar{\Xi}(y).
\]

We first observe that the extensive margin elasticities are non-negligible: of the order of 0.1 to 0.3 in absolute value (compare with the range of values obtained in the meta-analysis of Chetty, Guren, Manoli and Weber (2011): 0.17-0.26 for the steady-state extensive margin participation elasticities, and 0.33 for the intensive margin elasticities.

Figure 5: Extensive margin elasticities: \( \varepsilon = 0.33 \) and \( \varepsilon = 1 \)
The terms $A(y)$ in equation (61) are an order of magnitude smaller than the extensive margin elasticities.

The right hand side of formula (61) gives the social welfare effects of locally reforming the rate of progressivity of any, potentially suboptimal, tax schedule, e.g., the U.S. tax code (see Golosov, Tsyvinski and Werquin (2014)). The advantage of using a tax reform rather than an optimal tax approach for quantitative purposes is that the endogenous variables that appear in the formula (elasticities, income distributions, etc.) are all evaluated empirically given the actual U.S. data, without the need to extrapolate their values at the optimum tax schedule. This is strictly speaking what the “sufficient statistic” approach of Chetty (2009) allows us to do. Thus I compute the right hand side of (61) using the calibration described in Section 6.1, where the marginal value of public funds is given by the right hand side of equation (59) evaluated at the current U.S. tax code. I interpret the numerical results as the social welfare effects of raising the progressivity of the current U.S. tax code so that $1$ of statutory revenue would be mechanically raised absent any behavior changes.

I first show in Figure 6 the revenue effects of the tax reform disaggregated by income (unweighted by the density $f_y$), that is,

$$-T'(y) \frac{y \partial_p(y)}{1 - T'(y)} \varepsilon^*(y) \frac{1}{\mathbb{E} \partial_p} + T(y) \left[ \sum_{i=1}^{3} \frac{d \ln |\delta_i|}{dp} \Xi_i(y) \right] \frac{1}{\mathbb{E} \partial_p}.$$ 

The left and right panels plot these revenue effects by income in the frictionless and the frictional models for $\varepsilon = 0.33$ and $\varepsilon = 1$, respectively. The schedules of revenue effects are nearly identical at every income level. There are two reasons for getting such a small effect even though the extensive margin elasticities are non-negligible. First, note that the elasticities represented in Figure 5 have an opposite sign, so that the extensive margin terms partially cancel each other out in formula (61). Second, and most importantly, these elasticities are bounded, because a given change in progressivity changes the size of every individual’s inaction region and the volatility of their income process in the same proportion; in fact, the elasticities are roughly constant at the tails, where the income distribution is approximately Pareto distributed. But on the other hand, the increase in progressivity induces a much larger effect on the intensive margin (standard elasticity $\varepsilon^*$), because it increases the marginal tax rates that individuals face by an amount proportional to the log-income $\partial_p(y)$, which is unbounded. Therefore the option value effect on tax revenue is dominated by the standard intensive margin effect, more so for larger values of $\varepsilon$. 

37
I then show in Figure 7 the welfare effects of increasing progressivity disaggregated by (un-weighted) income, that is, I compute

\[- \left[ \partial_p (y) \omega (y) + \frac{d \ln \rho}{dp} \omega (y) \right] \frac{1}{\mathbb{E} \partial_p} + \left[ \sum_{i=1}^{3} \frac{d \ln |\delta_i|}{dp} \Omega_i (y) + \tilde{\Omega} (y) \right] \frac{1}{\mathbb{E} \partial_p}.\]

The left and right panels show the welfare effects in the frictionless and the frictional models for $\varepsilon = 0.33$ and $\varepsilon = 1$, respectively. These effects are large for the smaller value of the labor income elasticity $\varepsilon = 0.33$, and almost zero for the larger value $\varepsilon = 1$. The change in the distribution of utilities within incomes in response to a change in progressivity (captured by the welfare weight $\tilde{\Omega}$) plays little role in the discrepancy between the two curves in the left panel. This graphs show that for small enough intensive margin labor income elasticities, ignoring the hours restrictions within firms in the labor supply decisions of individuals leads to substantially mis-estimating the welfare costs of raising the progressivity of the tax schedule, by not taking into account the extensive margin effects. These effects disappear as the labor income elasticity gets higher, in which case the standard intensive margin welfare costs dominate the extensive margin effects. I discussed the intuition for the fact that the extensive margin effects on welfare tend to reduce the gains of raising progressivity in Section 5: an increase in progressivity is equivalent to a wider dispersion of individual incomes around their desired values, relative to the frictionless benchmark where the option value of adjusting behavior is unaffected by taxes. This negatively affects welfare.
I finally sum these effects over the whole population of incomes to obtain the integrals in the right hand side of (61), i.e., the net revenue and welfare effects, expressed in dollars, of a $1 statutory increase in tax revenue through an increase in the rate of progressivity. For $\varepsilon = 0.33$, in the frictionless model the total behavioral (revenue) loss of a $1$ statutory increase in $p$ is $c_{11.10}$, and the total welfare loss is $c_{83.12}$. The frictional effect of perturbing progressivity on revenue is 2.25% away from this frictionless effect (the behavioral loss is $c_{0.25}$ higher), and the frictional effect on welfare is 7.3% away from the frictionless effect (the welfare loss is $6.05$ higher). Thus wrongly assuming that the economy is frictionless leads to welfare calculations that underestimate the welfare costs of raising progressivity; this in turn will translate into a lower optimal rate of progressivity relative to the benchmark static model. In contrast, when $\varepsilon = 1$, the total behavioral revenue response of an increase in $p$ is $c_{33.44}$, and the total welfare loss is $c_{66.80}$. The frictional effects of perturbing progressivity are 0.46% and 0.75% away from the frictionless effects on revenue and welfare, respectively; therefore the static model's calculations are very accurate in this case. This is because when the labor income elasticity is high, most of the revenue and welfare effects are driven by the intensive margin of adjustment, which dwarves the extensive margin effects.

7 Conclusion

This paper analyzes a model where individual labor supply is subject to a fixed adjustment cost. The model allows for analytically tractable characterizations of the optimal individual behavior and the long-run aggregate income distributions in the presence of stochastic idiosyncratic wage shocks and a non-linear tax schedule. I derive a theoretical formula for the optimal long-run progressive

\footnotesize

\begin{itemize}
  \item Note that the sum of the two is lower than the mechanical effect $1$, implying that the U.S. tax code is too regressive for the parameters of the calibration. The optimum rate of progressivity is increasing the risk aversion $\gamma$ and decreasing in the elasticity $\varepsilon$.
  \item Note that the sum of the two is slightly larger than the mechanical effect $1$, implying that the U.S. tax code is slightly too progressive for this higher value of the labor income elasticity.
\end{itemize}
tax schedule in this frictional economy. I uncover several new effects that are not captured by standard frictionless optimal tax formulas with labor supply responses on the intensive margin. Most importantly, the option value of adjusting hours of work creates endogenously an extensive margin of labor supply conditional on participation which is affected by taxes, leading to novel welfare effects of non-linear tax instruments. The optimal tax schedule therefore depends on several new elasticities and marginal social welfare weights.

There are several directions for further research that would be interesting to study. First, it would be valuable to estimate numerically the novel effects highlighted in the theoretical formulas of this paper in a more sophisticated structural model. It would be valuable to allow for such features as savings and borrowings, life-cycle and endogenous participation labor supply choices, non-proportional fixed adjustment costs, transitory as well as permanent wage shocks, income-varying labor supply elasticities, more general non-linear taxes and transfer programs, and other dimensions of labor supply adjustment choices (e.g., job satisfaction). Such a model would also provide realistic estimates of the speed of adjustment of the economy in response to tax changes, and hence allow for a comparison of the (empirically observed) short-run elasticities versus the (inferred) long-run elasticities in the presence of sluggish responses of individuals to taxes. It would also be interesting to add aggregate shocks to the model and study the effects of taxes over the business cycle as aggregate shocks push the densities of incomes towards or away from the boundaries of the inaction region. On the theoretical side, it would be worthwhile to characterize the optimal tax schedule in environments with fixed costs of adjustment both on the labor supply and the labor demand with an endogenous wage, leading to bilateral monopoly situations. I leave these important questions for future research.

References


A Appendix

Individual behavior and aggregation

I first derive several results about individual welfare:

**Proof of Proposition 1 and equations (33), (34), (53).** Substituting $y^*$ for $w$ using the optimality condition (10) into the flow utility $U(y - T(y), \frac{y}{w})$ yields

$$U = \frac{1}{1 - \gamma} \left[ \frac{1 - \tau}{1 - p} (y^*)^{1-p} \right]^{1-\gamma} \times \left[ \left( \frac{y}{y^*} \right)^{1-p} - \frac{1 - p}{1 + 1/\varepsilon} \left( \frac{y}{y^*} \right)^{1+1/\varepsilon} \right]^{1-\gamma},$$

which, using $y = y^* e^{\delta}$, implies (15). A second-order Taylor approximation of the function $u(\delta)$ around the frictionless optimum $\delta = 0$ easily yields equation (16). The geometric Brownian motion process (GBM) for $c^*$ implies

$$dg(c^*_t) = \left[ (1 - \gamma) \mu_c + \frac{1}{2} (1 - \gamma)^2 \sigma_c^2 \right] g(c^*_t) dt + \left[ (1 - \gamma) \sigma_c \right] g(c^*_t) d\mathcal{W}_t,$$

and hence $g(c^*_t) = g(c^*_0) e^{(1-\gamma)\mu_c t + (1-\gamma)\sigma_c \mathcal{W}_t}$. In the frictionless model, the individual value function is then given by

$$\mathcal{V}(y) = E \left[ \int_0^\infty e^{-(\rho_1+\rho_2)t} \frac{1}{1 - \gamma} \left( c^*_t - \frac{1}{1 + 1/\varepsilon} \left( \frac{y^*_t}{w_t} \right)^{1+1/\varepsilon} \right)^{1-\gamma} dt \mid y = y^*(w_0) \right]$$

$$= \frac{(1+p_\varepsilon)}{1 + \gamma} \mathcal{E}_0 \left[ \int_0^\infty e^{-(\rho_1+\rho_2)t} (c^*_t)^{1-\gamma} dt \right] = \frac{(1+p_\varepsilon)}{1 + \gamma} \int_0^\infty e^{-(\rho_1+\rho_2+1-\gamma)\mu_c t - \frac{1}{2} (1-\gamma)^2 \sigma_c^2 t},$$

where the last equality follows from the facts that $\mathcal{W}_t \sim \mathcal{N}(0, t)$ and $E[e^{X}] = e^{\mu + \sigma^2/2}$ if $X \sim \mathcal{N}(\mu, \sigma^2)$. Equation (33) follows. To compute the marginal social welfare weights (53), let $\hat{c}_t$ follow the same GBM process as $c^*_t$, implying that $\frac{1+p_\varepsilon}{1 + \varepsilon} c^*_t + \hat{c}_t$ also follows the same process. Hence the same steps as above imply

$$\omega^*(y) = \frac{1}{\lambda^*} \frac{d}{dc_0} \left( \frac{1}{1 - \gamma} \left( \frac{1+p_\varepsilon}{1 + \varepsilon} c^*_0 + \hat{c}_0 \right)^{1-\gamma} \frac{1}{\rho_1 + \rho_2 - (1 - \gamma) \mu_c - \frac{1}{2} (1 - \gamma)^2 \sigma_c^2} \right)_{\hat{c}_0=0},$$

leading to (53).

In the frictional model, the value function for inactive individuals is equal to

$$\mathcal{V}_i(y^*_0, \delta_0) = \max_{\tau_1} \mathcal{E}_0 \left[ \int_0^{\tau_1} e^{-(\rho_1+\rho_2)t} g \left( \frac{1 - \tau}{1 - p} y_t^{1-p} \right) u(\delta_t) dt + e^{-(\rho_1+\rho_2)\tau_1} \left\{ \mathcal{V}_s(y^*_{\tau_1}, \delta_{\tau_1}) - \kappa_{\tau_1} \right\} \right],$$

subject to the laws of motion (11) and (14), where $\tau_1$ is the optimal stopping time at which the
individual starts searching by paying the fixed cost, and the value function of searchers is equal to

\[ V_s(y^*_i, \delta_0) = \max_{\delta^*_t} \mathbb{E}_0 \left[ \int_0^{\tau_s} e^{-(\rho_1+\rho_2)t} U_y(y^*_i, \delta_t) \, dt + e^{-(\rho_1+\rho_2)\tilde{\tau}} V_i(y^*_i, \delta^*_\tilde{\tau}) \right], \]

subject to the laws of motion (11) and (14), where \( \tilde{\tau} \) is a stopping time with a Poisson an exponential distribution with parameter \( q \), and \( \delta^*_\tilde{\tau} \) is the optimal deviation that the individual chooses upon reception of an adjustment opportunity at time \( \tilde{\tau} \). From these equations we can write a sequential formulation of the value functions \( V_i \) and \( V_s \) and obtain that

\[ \frac{V_i(y^*, \delta)}{g\left(\frac{V_i}{1+\lambda y^{x-1}}\right)} = v_x(\delta) \quad \text{for} \quad x \in \{i, s\}, \]

where

\[
\begin{align*}
v_i(\delta_0) &= \max_{\delta^*_t} \mathbb{E}_0 \left[ \int_0^{\tau_i} e^{-(\rho_1+\rho_2-(1-\gamma)\mu_\sigma)t+(1-\gamma)\sigma \mathcal{W}_i \, u(\delta_t) \, dt ight] \\
&\quad + e^{-(\rho_1+\rho_2-(1-\gamma)\mu_\sigma)\tau_i+(1-\gamma)\sigma \mathcal{W}_i \{ v_s(\delta_{\tau_i}) - \kappa \}} \\
v_s(\delta_0) &= \max_{\delta^*_s} \mathbb{E}_0 \left[ \int_0^{\tau_s} e^{-(\rho_1+\rho_2-(1-\gamma)\mu_\sigma)t+(1-\gamma)\sigma \mathcal{W}_s \, u(\delta_t) \, dt ight] \\
&\quad + e^{-(\rho_1+\rho_2-(1-\gamma)\mu_\sigma)\tilde{\tau}+(1-\gamma)\sigma \mathcal{W}_\tilde{\tau} v_i(\delta^*_\tilde{\tau})} \]
\end{align*}
\]

Proposition 1 and equation (34) follow.

I next prove the Verification Proposition 2:

**Proof of Proposition 2.** Fix an impulse control policy \( p = \left\{ (\tau_j, \Delta_j (\tilde{\tau}_j)) \right\}_{j \geq 1} \in \mathcal{P} \). By condition (i) and Theorem 2.1. in Oksendal and Sulem (2007), we can assume that \( v \in \mathcal{C}^2(\mathbb{R}) \). Hence using condition (iv) we can apply the localized version of Dynkin’s formula (Theorem 1.24. in Oksendal and Sulem (2007) modified to take into account the discounting) to get, for \( j \geq 0, \)

\[
\mathbb{E} \left[ e^{-(\rho_1+\rho_2)(\tau_j+\tilde{\tau}_j)} g \left( c^*_j, \tilde{\tau}_j \right) \right] - \mathbb{E} \left[ e^{-(\rho_1+\rho_2)\tau_{j+1}} g \left( c^*_j, \tilde{\tau}_{j+1} \right) \right] = - g \left( c^*_0 \right) \mathbb{E} \left[ \int_{\tau_j+\tilde{\tau}_j}^{\tau_{j+1}} e^{-(\rho_1+\rho_2-(1-\gamma)\mu_\sigma)t+(1-\gamma)\sigma \mathcal{W}_i (-pv(\delta_t) + \mathcal{L}v(\delta_t)) \, dt \right] \\
\geq g \left( c^*_0 \right) \mathbb{E} \left[ \int_{\tau_j+\tilde{\tau}_j}^{\tau_{j+1}} e^{-(\rho_1+\rho_2-(1-\gamma)\mu_\sigma)t+(1-\gamma)\sigma \mathcal{W}_s u(\delta_t) \, dt \right],
\]

where the inequality follows from condition (22), and it becomes an equality if \( p = p^* \). Moreover,
we have
\[
\mathbb{E} \left[ e^{-(\rho_1+\rho_2)\tau_{j+1}} g \left( c^*_{\tau_{j+1}} \right) v \left( \delta_{\tau_{j+1}} \right) \right] \\
- \mathbb{E} \left[ e^{-(\rho_1+\rho_2)(\tau_{j+1}+\tau_j)} g \left( c^*_{\tau_{j+1}+\tau_j} \right) v \left( \delta_{\tau_{j+1}+\tau_j} + \Delta_{j+1}(\tau_{j+1}) \right) \right] \\
\geq g (c^*_0) \left\{ \mathbb{E} \left[ e^{-(\rho_1+\rho_2-(1-\gamma)\mu_c)\tau_{j+1}+(1-\gamma)\sigma_c W_{\tau_{j+1}} v \left( \delta_{\tau_{j+1}} \right) \right] \right. \\
- \mathbb{E} \left[ e^{-(\rho_1+\rho_2-(1-\gamma)\mu_c)(\tau_{j+1}+\tau_j)+(1-\gamma)\sigma_c W_{\tau_{j+1}+\tau_j}} \sup_{\delta^* \in \mathbb{R}} v \left( \delta^* \right) \right] \right\} \\
= g (c^*_0) \mathbb{E} \left[ e^{-(\rho_1+\rho_2-(1-\gamma)\mu_c)\tau_{j+1}+(1-\gamma)\sigma_c W_{\tau_{j+1}}} \left( v \left( \delta_{\tau_{j+1}} \right) - \mathcal{M} v \left( \delta_{\tau_{j+1}} \right) + \hat{v}_s (\delta_{\tau_{j+1}} - \kappa) \right) \right] \\
\geq g (c^*_0) \mathbb{E} \left[ e^{-(\rho_1+\rho_2-(1-\gamma)\mu_c)\tau_{j+1}+(1-\gamma)\sigma_c W_{\tau_{j+1}}} \left( \hat{v}_s (\delta_{\tau_{j+1}} + \kappa) \right) \right],
\]
where the last equality follows from condition (ii). Both inequalities become equalities if \( p = p^* \).
Thus, we obtain, summing the previous equations from \( j = 0 \) to \( j = N \geq 1 \),
\[
g (c^*_0) v (\delta_0) - \mathbb{E} \left[ e^{-(\rho_1+\rho_2)(\tau_{N+1}+\tau_{N+1})} g \left( c^*_{\tau_{N+1}+\tau_{N+1}} \right) v \left( \delta_{\tau_{N+1}} \right) \right] \\
\geq g (c^*_0) \mathbb{E} \left[ \sum_{j=0}^{N} \int_{\tau_{j+1}}^{\tau_{j+1}+\tau_j} e^{-(\rho_1+\rho_2-(1-\gamma)\mu_c)\tau_{j+1}+(1-\gamma)\sigma_c W_{\tau_{j+1}}} u (\delta_t) dt \\
+ \sum_{j=0}^{N} e^{-(\rho_1+\rho_2-(1-\gamma)\mu_c)\tau_{j+1}+(1-\gamma)\sigma_c W_{\tau_{j+1}}} \left( \hat{v}_s (\delta_{\tau_{j+1}} - \kappa) \right) \right] \\
= \mathbb{E} \left[ \int_{0}^{\tau_{N+1}+\tau_{N+1}} e^{-(\rho_1+\rho_2)t} g (c^*_t) u (\delta_t^p) dt - \sum_{j=0}^{N} e^{-(\rho_1+\rho_2)\tau_{j+1}} \kappa g \left( c^*_{\tau_{j+1}} \right) \right],
\]
(with equality if \( p = p^* \)), where the equality follows from (19). Now, as \( N \to \infty \), we have \( \tau_N \to \infty \) so that the second term on the l.h.s. of the previous equation converges to zero. Therefore we obtain
\[
\mathcal{V} \left( y^*_0, \delta_0 \right) \geq \mathbb{E} \left[ \int_{0}^{\infty} e^{-(\rho_1+\rho_2)t} U \left( y^*_t, \delta_t^p \right) dt - \sum_{j=0}^{\infty} e^{-(\rho_1+\rho_2)\tau_{j+1}} \kappa g \left( c^*_{\tau_{j+1}} \right) \right], \forall p \in \mathcal{P}
\]
with equality if \( p = p^* \). (We restrict the set \( \mathcal{P} \) of admissible controls to those \( p \) that satisfy \( \mathbb{E} \left[ \int_0^{\infty} e^{-(\rho_1+\rho_2)t} \left| u \left( \delta_t^p \right) \right| dt \right] < \infty \).) But the right hand side is the lifetime utility of an individual at birth, after \( \delta_0 \) has been chosen, under the control policy \( p \) (see (7)). This concludes the proof.

The next proof shows that in the special case where \( q = \infty \), the optimal individual policy is characterized by a single inaction band \( (\hat{\delta}, \delta^*, \delta) \).

**Proof of Proposition 1.** To prove that the conjectured policy is indeed optimal, we need to show that if the value function satisfies (24,25,26,27,28), then it satisfies the assumptions of the Verification Proposition 2 and the quasi-variational inequalities (21,22). If \( q = \infty \), then the technical
conditions required for the Verification theorem 2 to hold are simpler (see Richard (1977)): \( v' \) must be absolutely continuous and bounded and \( v'' \) must be in \( L^2(\mathbb{R}) \); these are easily verified. It remains to check that the QVI are satisfied.

First, I show that the lower bound of the conjectured inaction region is non-positive, \( \hat{\delta} \leq 0 \), and the upper bound is non-negative, \( \bar{\delta} \geq 0 \), so that the argmax of the flow utility \( u(\delta) \) lies within the inaction region. Note first that \( v''(\hat{\delta}^+) > 0, v''(\hat{\delta}^-) < 0 \) and \( v''(\delta^*) \leq 0 \), where the first two inequalities follow from the continuity of \( v' \) and the assumption that \( v' > 0 \) (resp., \( v' < 0 \)) in a neighborhood to the right of \( \hat{\delta} \) (resp. to the left of \( \bar{\delta} \)). Define the first and the last inflection points of \( v \) on \( [\hat{\delta}, \delta^*] \) by We know that such values exist in \([\hat{\delta}, \delta^*]\), because \( v''(\hat{\delta}^+) > 0 \) and \( v'(\delta^*) = 0 \). Since \( u, v, v' \in C^1((\hat{\delta}, \delta^*)) \), the HJB equation (24) implies that \( v'' \) is continuously differentiable. Taking left derivatives of the HJB and evaluating at \( \delta^*_M \) yields

\[
\frac{\sigma_c^2}{2} v''(\delta^*_M) = \rho v'(\delta^*_M) - u'(\delta^*_M) \geq \rho v'(\delta^*) - u'(\delta^*_M) = -u'(\delta^*_M),
\]

where the inequality follows from the definition of \( \delta^*_M \) and \( v''(\delta^*) \leq 0 \). Since \( v''(\delta^*_M) \leq 0 \), we obtain \(-u'(\delta^*_M) \leq 0 \). But \( u \) is concave with a unique global maximum at \( \delta = 0 \), hence \( \delta^*_M \leq 0 \) and \( u'(\delta) > u'(\delta^*_M) \) for all \( \delta < \delta^*_M \).

Second, I show that the conjectured value function \( v \) is unimodal, that is, \( v(\delta) \) is strictly increasing on \((\hat{\delta}, \delta^*)\) and strictly decreasing on \((\delta^*, \bar{\delta})\). If \( \delta^*_m = \delta^*_M \), then \( v'(\delta) > 0 \) for all \( \delta \in (\hat{\delta}, \delta^*) \). Suppose that \( \delta^*_m < \delta^*_M \) and there exists \( \delta \in (\hat{\delta}, \delta^*) \) such that \( v'(\delta) \leq 0 \). Then there exists \( \delta \in (\delta^*_m, \delta^*_M) \) which is a local minimizer of \( v' \), with \( u'(\delta) < 0 \) and \( v''(\delta^-) \geq 0 \). Taking left derivatives in the HJB equation (24) and evaluating at \( \delta^- \) yields:

\[
u'(\delta^-) = -\frac{1}{2} \sigma_c^2 (\delta^-)^2 v''(\delta^-) + \rho v'(\delta^-) < 0.
\]

But we saw above that \( u'(\delta^*_M) \geq 0 \), which implies \( u'(\delta^-) > 0 \) since \( \delta^- \delta^*_M \) and \( u \) is strictly concave and unimodal, a contradiction.

I now show that the conjectured value function satisfies the QVI (22), i.e., \( \mathcal{L} v(\delta) - \rho v(\delta) + u(\delta) \leq 0 \) for all \( \delta \in \mathbb{R} \). The HJB equation (24) and a symmetry argument imply that it is sufficient to check this inequality on \((\infty, \hat{\delta})\). Fix \( \delta \in \hat{\delta} \). We have

\[
[\mathcal{L} v(\delta) - \rho v(\delta) + u(\delta)] = [\mathcal{L} v(\delta^+) - \rho v(\delta) + u(\delta)] - [\rho (v(\delta) - v(\delta))] - [\mu_y + (1 - \gamma) \sigma_c \sigma_y] v'(\delta) - v'(\delta) = \frac{1}{2} \sigma_c^2 \sigma^2 (v''(\delta) - v''(\delta^*)) + (u(\delta) - u(\delta)).
\]

The HJB equation (24), the value-matching condition (25) and the smooth-pasting conditions (26) imply that the first line of the right hand side is equal to zero. Moreover, we saw above that \( v''(\delta) = 0 < v''(\delta^+) \) and \( u(\delta) \leq u(\delta) \), which concludes the proof.

Finally, I show that the conjectured value function satisfies the QVI (21), i.e., \( \mathcal{M} v(\delta) - v(\delta) \leq 0 \) for all \( \delta \in \mathbb{R} \), where \( \mathcal{M} v(\delta) = v(\delta^*) - \kappa \). For \( \delta \in \mathbb{R} \setminus (\hat{\delta}, \bar{\delta}) \), this QVI is satisfied with equality by
Integrating over the stopping time relevant boundary conditions at expression. (Formula 28 is obtained by solving directly the HJB equation for searchers with the Straightforward algebra finally shows that the first term of equation (28) is equal to the previous $$\delta$$ letting $$\hat{\delta}$$.

Proof of equation (28). The value of searchers can be calculated explicitly as follows. We have, letting $$\hat{\xi} \sim \mathcal{E}(q),$$

$$\mathbb{E}_0 \left[ \int_0^{\hat{\tau}} e^{-(\rho_1+\rho_2)t} g(c_t^*) u(\delta_t) \, dt + e^{-(\rho_1+\rho_2)\hat{\tau}} g(c_{\hat{\tau}}^*) v_i(\delta^*) \right]$$

$$= g(c_0^*) \int_0^\infty q e^{-qt} \mathbb{E}_0 \left[ \int_0^t e^{-(\rho_1+\rho_2-(1-\gamma)\mu_c)s+(1-\gamma)\sigma_c W_s} u(\delta_s) \, ds \right.$$

$$+ e^{-(\rho_1+\rho_2-(1-\gamma)\mu_c)\hat{\tau}+(1-\gamma)\sigma_c W_{\hat{\tau}}} v_i(\delta^*) \right] \, dt.$$

Using Fubini’s theorem, the fact that $$W_t \sim \mathcal{N}(0, t),$$ and the quadratic approximation $$u(\delta) = \alpha_0 + \alpha_1\delta + \alpha_2\delta^2,$$ we obtain that the stochastic integral (for fixed $$t$$) is equal to, with $$\delta_s = \delta_0 + \mu_\delta s + \sigma_\delta W_s$$:

$$\mathbb{E}_0 \left[ \int_0^t e^{-(\rho_1+\rho_2-(1-\gamma)\mu_c)s+(1-\gamma)\sigma_c W_s} u(\delta_s) \, ds \right]$$

$$= \frac{1}{\rho^2} \alpha_2 \left( (1-\gamma) \sigma_c \sigma_\delta + \mu_\delta \right)^2 \left( 2 - e^{-\rho t} \left( \rho^2 t^2 + 2\rho t + 2 \right) \right) + \frac{1}{\rho} \left( \alpha_0 + \alpha_1 \delta_0 + \alpha_2 \delta_0^2 \right) (1 - e^{-\rho t})$$

$$+ \frac{1}{\rho^2} \left( \alpha_1 (\mu_\delta + (1-\gamma) \sigma_c \sigma_\delta) + \alpha_2 \sigma_\delta^2 + 2\alpha_2 (\mu_\delta + (1-\gamma) \sigma_c \sigma_\delta) \delta_0 \right) (1 - e^{-\rho t} (\rho t + 1)).$$

Integrating over the stopping time $$t$$ then yields:

$$\int_0^\infty q e^{-qt} \mathbb{E}_0 \left[ \int_0^t e^{-(\rho_1+\rho_2-(1-\gamma)\mu_c)s+(1-\gamma)\sigma_c W_s} u(\delta_s) \, ds \right] \, dt$$

$$= \left( \frac{\alpha_2}{\rho + q} \right) \sigma_\delta^2 + \left( \frac{\alpha_1}{\rho + q} + 2\alpha_2 \frac{\mu_\delta + (1-\gamma) \sigma_c \sigma_\delta}{(\rho + q)^2} \right) \delta_0$$

$$+ \left( \frac{\alpha_0}{\rho + q} + \alpha_1 \frac{(\mu_\delta + (1-\gamma) \sigma_c \sigma_\delta)}{(\rho + q)^2} + \alpha_2 \frac{2(\mu_\delta + (1-\gamma) \sigma_c \sigma_\delta)^2 + (\rho + q) \sigma_\delta^2}{(\rho + q)^3} \right).$$

Straightforward algebra finally shows that the first term of equation (28) is equal to the previous expression. (Formula 28 is obtained by solving directly the HJB equation for searchers with the relevant boundary conditions at $$\pm \infty.$$) Finally the value of returning to $$\delta^*$$ is given by

$$\int_0^\infty q e^{-qt} \mathbb{E}_0 \left[ e^{-(\rho_1+\rho_2-(1-\gamma)\mu_c)t+(1-\gamma)\sigma_c W_t} v_i(\delta^*) \right] \, dt = \frac{qv_i(\delta^*)}{\rho_1 + \rho_2 + q - (1-\gamma) \mu_c - \frac{1}{2} (1-\gamma)^2 \sigma_c^2},$$

which is the second term of equation (28).

I now provide an intuitive derivation of the QVI of Proposition 2.

50
**Heuristic derivation of equations (21) and (22).** Suppose that an optimal policy $p^*$ exists. If the individual adopts an arbitrary control for an infinitesimal amount of time and then switches back to the optimal control $p^*$, then the resulting value function cannot be better than the optimal one. If the individual is currently inactive, there are only two possible choices of control during that infinitesimal period: not impose any control (QVI (22)), and pay the fixed cost to begin a search period (QVI (21)). Finally, one of the two quasi-variational inequalities must hold with equality, since one of these two choices of control must be optimal.

Consider an inactive individual with frictionless income and deviation $(y^*_i, \delta_i)$, who remains inactive during the time interval $[t, t + \Delta t)$ for some small $\Delta t > 0$, then reverts back to the optimal policy $p^*$. Her value function $V(y^*_i, \delta_i)$ satisfies

$$V(y^*_i, \delta_i) \geq U(y^*_i, \delta_i) \Delta t + \frac{1 - \rho_2 \Delta t}{1 + \rho_1 \Delta t} \mathbb{E}_t [V(y^*_i + \Delta y^*, \delta_i + \Delta \delta)],$$

where $\Delta \delta = -(1 - \frac{\rho_1}{1 + \varepsilon}) \Delta \ln y^*$. Multiplying by $(1 + \rho_1 \Delta t)$, subtracting $(1 - \rho_2 \Delta t) V(y^*_i, \delta_i)$ and dividing by $\Delta t$ on both sides, we obtain, letting $\Delta t \to 0$,

$$(\rho_1 + \rho_2) V(y^*, \delta) \geq U(y^*, \delta) + \frac{\mathbb{E}_t [dV(y^*, \delta)]}{dt}.$$

Using Itô’s formula and the laws of motion of $y^*_t$ and $\delta_t$, we find

$$\mathbb{E}_t [dV(y^*_t, \delta_t)] = \left( \mu_y + \frac{1}{2} \sigma_y^2 \right) y^*_t \frac{\partial V}{\partial y^*} + \mu_\delta \frac{\partial V}{\partial \delta} + \frac{1}{2} \sigma_y^2 (y^*_t)^2 \frac{\partial^2 V}{\partial (y^*)^2} + \frac{1}{2} \sigma_\delta^2 \frac{\partial^2 V}{\partial \delta^2} + \sigma_y \sigma_\delta y^*_t \frac{\partial^2 V}{\partial y^* \partial \delta}. $$

Since the value function is homogeneous in $g(c^*)$, we can replace $V(y^*_t, \delta_t)$ with $g(c^*_t) v(\delta_t)$ in the resulting equation and divide through by $g(c^*_t)$ to obtain $v(\delta_t) \geq L v(\delta_t) + u(\delta_t)$.

Next suppose that the individual pays the fixed adjustment cost at time $t$, and hence becomes a searcher. We have

$$V(y^*_t, \delta_t) \geq V_s(y^*_t, \delta_t) - \kappa g(c^*_t).$$

Dividing both sides by $g(c^*_t)$ and using the expression derived above for $v_s(\delta)$, we thus obtain $v(\delta_t) \geq M v(\delta_t)$. \hfill \Box

Feng and Muthuraman (2010) provide an algorithm to compute numerically the optimal individual policy solution to (24, 25, 25, 27), which is easily extended to this paper’s environment.

Next I derive the characterization of the stationary income and deviation distributions.

**Proof of Proposition 4.** To derive the KFE equations (39) that must be satisfied by the stationary joint distributions $f^i_{\ln y^*, \delta}$ and $f^s_{\ln y^*, \delta}$, discretize the processes $(\ln y^*, \delta)$ on a two-dimensional grid with size $\left(\Delta h, \frac{1 - \rho_1}{1 + \varepsilon} \Delta h\right)$. In the time unit $\Delta t$, $\ln y^*$ moves up by $\Delta h = \sigma_y \sqrt{\Delta t}$ and $\delta$ moves down by $\frac{1 - \rho_1}{1 + \varepsilon} \Delta h$ with probability $\frac{1}{2} \left(1 + \frac{\rho_1}{\sigma_y^2} \Delta h\right)$. The balanced flow equations for $f^i_{\ln y^*, \delta}$ at point
\[(u, \delta) \in \mathbb{R} \times \{(\tilde{\delta}, \delta^*) \cup (\delta^*, \tilde{\delta})\}\]

write:

\[
f_{ln,y^*,\delta}^i (u, \delta) = (1 - \rho_2 \Delta t) \left\{ \frac{1}{2} \left( 1 + \frac{\mu_y}{\sigma_y^2} \Delta h \right) f_{ln,y^*,\delta}^i \left( u - \Delta h, \delta + \frac{1 - p}{1 + 1/\varepsilon} \Delta h \right) \right. \\
+ \frac{1}{2} \left( 1 - \frac{\mu_y}{\sigma_y^2} \Delta h \right) f_{ln,y^*,\delta}^i \left( u + \Delta h, \delta - \frac{1 - p}{1 + 1/\varepsilon} \Delta h \right) \right\}.
\]

Taking a second-order Taylor expansion in \(\Delta h\) of this equation around 0 and rearranging terms easily yields (39). We also have \(f_{ln,y^*,\delta}^i = 0\) if \(\delta \notin (\tilde{\delta}, \tilde{\delta})\). The balanced flow equations for \(f_{ln,y^*,\delta}^s\) at point \((u, \delta) \in \mathbb{R} \times \{(-\infty, \tilde{\delta})\cup(\tilde{\delta}, \delta^*)\cup(\delta^*, \infty)\}\) write identically, except that the right hand side is multiplied by the probability \((1 - q\Delta t)\) of exiting the search area in \([t, t + \Delta t]\) due to the arrival of an adjustment opportunity.

The boundary conditions (41) to (45) can be derived as follows. The balanced-flow equation for \(f_{ln,y^*,\delta}^i\) at the point \((u, \tilde{\delta})\) writes:

\[
f_{ln,y^*,\delta}^i (u, \tilde{\delta}) = (1 - \rho_2 \Delta t) \left\{ \frac{1}{2} \left( 1 - \frac{\mu_y}{\sigma_y^2} \Delta h \right) f_{ln,y^*,\delta}^i \left( u + \Delta h, \tilde{\delta} - \frac{1 - p}{1 + 1/\varepsilon} \Delta h \right) \right\}.
\]

A first-order Taylor expansion in \(\Delta h\) around 0 yields \(f_{ln,y^*,\delta}^i (u, \tilde{\delta}) = 0\), and similarly \(f_{ln,y^*,\delta}^s (u, \tilde{\delta}) = 0\). Similarly, the balanced flow condition at the boundaries \(\pm\infty\) for \(f_{ln,y^*,\delta}^s\) writes

\[
\lim_{\delta \to \pm\infty} f_{ln,y^*,\delta}^s (u, \delta)|_h \text{ constant} = 0.
\]

Noting that the condition “\(h\) constant” is equivalent to “\((1-p)/1+\varepsilon \ln y^* + \delta \) constant”, we obtain (42). The balanced-flow equation for \(f_{ln,y^*,\delta}^i\) at the point \((u, \delta^*)\) writes:

\[
f^i (u, \delta^*) = (1 - \rho_2 \Delta t) \left\{ \frac{1}{2} \left( 1 + \frac{\mu_y}{\sigma_y^2} \Delta h \right) f^i \left( u - \Delta h, \delta^* + \frac{1 - p}{1 + 1/\varepsilon} \Delta h \right) \right. \\
+ \frac{1}{2} \left( 1 - \frac{\mu_y}{\sigma_y^2} \Delta h \right) f^i \left( u + \Delta h, \delta^* - \frac{1 - p}{1 + 1/\varepsilon} \Delta h \right) \right\} \\
+ (1 - \rho_2 \Delta t) (q\Delta t) \left\{ \frac{1}{2} \left( 1 + \frac{\mu_y}{\sigma_y^2} \Delta h \right) \sum_{\delta \in G} f^s (u - \Delta h, \delta) \right. \\
+ \frac{1}{2} \left( 1 - \frac{\mu_y}{\sigma_y^2} \Delta h \right) \sum_{\delta \in G} f^s (u + \Delta h, \delta) \right\} \\
+ (\rho_2 \Delta t) \left( \frac{1 - p}{1 + 1/\varepsilon} \Delta h \right)^{-1} f_{ln,y^*_0} (u),
\]

where \(G\) denotes the grid of \(\delta\) and \(f_{ln,y^*_0}\) denotes the density of frictionless log-incomes at birth.
Taking a first-order Taylor expansion in $\Delta h$ around 0 using

$$\sum_{\delta \in G} f_{\ln y^*, \delta}^s (\ln y^*, \delta) \xrightarrow{\Delta h \to 0} \left( \frac{1 - p}{1 + 1/\varepsilon} \Delta h \right)^{-1} f_{\ln y^*}^s (\ln y^*)$$

yields (43). Finally, the balanced flow condition at the boundary $\bar{\delta}$ for $f_{\ln y^*, \delta}^s$ writes:

\[
\begin{align*}
    f^s(u, \bar{\delta}) &= (1 - \rho_2 \Delta t)(1 - q \Delta t) \left\{ \frac{1}{2} \left( 1 + \frac{\mu y}{\sigma^2_y} \Delta h \right) f^s(u - \Delta h, \bar{\delta} + \frac{1 - p}{1 + 1/\varepsilon} \Delta h) \\
    &+ \frac{1}{2} \left( 1 - \frac{\mu y}{\sigma^2_y} \Delta h \right) f^s(u + \Delta h, \bar{\delta} - \frac{1 - p}{1 + 1/\varepsilon} \Delta h) \right\} \\
    &+ (1 - \rho_2 \Delta t) \left\{ \frac{1}{2} \left( 1 - \frac{\mu y}{\sigma^2_y} \Delta h \right) f^i(u + \Delta h, \bar{\delta} - \frac{1 - p}{1 + 1/\varepsilon} \Delta h) \right\}.
\end{align*}
\]

A first-order Taylor expansion in $\Delta h$ yields (45). Equation (44) is obtained similarly.

Note that changing variables and defining the density

\[
g^i(u, \delta) = \frac{1 - p}{1 + 1/\varepsilon} f_{\ln y^*, \delta}^i \left( u - \frac{1 - p}{1 + 1/\varepsilon} \delta, \delta \right)
\]

implies that $g^i$ satisfies the PDE

\[
\frac{1}{2} \sigma^2_y g^i_{22} + \mu \delta_2 g^i_2 - \rho_2 g^i = 0 \quad \text{on} \ \mathbb{R} \times \{(\bar{\delta}, \delta^*) \cup (\delta^*, \bar{\delta})\}.
\]

Imposing the boundary conditions above on the explicit solution to this differential equation yields, after tedious algebra and letting $\Delta \equiv \delta^* - \delta$ and $\bar{\Delta} \equiv \bar{\delta} - \delta^*$,

\[
\begin{align*}
    f_{\ln y^*, \delta}^i(u, \delta) &= \tilde{f} \left( u + \frac{1 + 1/\varepsilon}{1 - p} \delta \right) \left\{ e^{-r_{1,\delta}^2 (\delta - \delta)} - e^{-r_{2,\delta}^2 (\delta - \bar{\delta})} \right\} (\delta^* - \delta) \\
    &= \tilde{f} \left( u + \frac{1 + 1/\varepsilon}{1 - p} \delta \right) \left\{ e^{-r_{1,\delta}^2 (\delta - \delta)} - e^{-r_{2,\delta}^2 (\delta - \bar{\delta})} \right\} (\delta^* - \delta)
\end{align*}
\]

where the function $\tilde{f} (\cdot)$ satisfies the normalization (46) and the integral equation

\[
\tilde{f}(u) = \left\{ \frac{e^{-r_{1,\delta}^2 \Delta} - e^{-r_{2,\delta}^2 \Delta}}{e^{r_{1,\delta}^2 \Delta} - e^{-r_{2,\delta}^2 \Delta}} \right\} \left\{ \frac{2 \rho_2}{\sigma^2_y} f_{\ln y_0^*} \left( u - \frac{1 + 1/\varepsilon}{1 - p} \delta \right) \right\} + \ldots
\]

\[
\ldots + \frac{2 q}{\sigma_\delta^2} \int_{-\infty}^{\infty} f_{\ln y^*, \delta}^s \left( u - \frac{1 + 1/\varepsilon}{1 - p} \delta^*, \delta \right) d\delta
\]

53
for all $u \in \mathbb{R}$. Now suppose $q = \infty$. The functional equation satisfied by $\tilde{f}$ is simpler:

$$\frac{2\rho_2}{\sigma_y^2 \delta} \frac{1+1/\varepsilon}{1-p} r_{1,\delta}^{\rho_2} \int \ln y_0^*(u) \, dy = \tilde{f} \left( u + \frac{1+1/\varepsilon \delta}{1-p} \right) - \tilde{f} \left( u + \frac{1+1/\varepsilon \delta}{1-p} \right) = \left( \frac{r_{2,\delta}^{\rho_2} e^{-r_{2,\delta}^{\rho_2}} - e^{-r_{1,\delta}^{\rho_2}}}{e^{-r_{2,\delta}^{\rho_2}} - e^{-r_{1,\delta}^{\rho_2}}} \right) - \left( \frac{r_{2,\delta}^{\rho_2} e^{-r_{2,\delta}^{\rho_2}} - e^{-r_{1,\delta}^{\rho_2}}}{e^{-r_{2,\delta}^{\rho_2}} - e^{-r_{1,\delta}^{\rho_2}}} \right) \tilde{f} \left( u + \frac{1+1/\varepsilon \delta}{1-p} \right).$$

Letting $\tilde{f} (u) = e^{1+1/\varepsilon \delta i \beta u} g (u)$, it is easy to check that the solution $g (u)$ to the equation converges to a constant as $u \to \infty$. This implies that

$$f_{\ln y^*}^i (u) = \int_{\delta}^{\hat{\beta}} f_{\ln y^*, \delta} (u, \delta) \, d\delta \sim e^{r_{1,\delta}^{\rho_2} u},$$

i.e., that the stationary income distribution has a Pareto right tail with coefficient $r_{1,\delta}^{\rho_2}$. 

**Optimal Taxation**

I first derive the formulas (59) and (60) characterizing the optimal taxes in the frictionless model.

**Proof of equations (59) and (60).** Consider a perturbation $(d\tau, dp)$ of the baseline tax system. The first-order change in the tax liability at income $y$ is, to a first-order in $(d\tau, dp)$ (ignoring the term $o (d\tau, dp)$),

$$\tilde{T} (y) - T (y) = \left( y - \frac{1-\tau - d\tau}{1-p - dp} y^{1-p- dp} \right) - \left( y - \frac{1-\tau}{1-p} y^{1-p} \right) = \frac{1}{1-p} y^{1-p} d\tau + \left( \ln y - \frac{1}{1-p} \right) \frac{1-\tau}{1-p} y^{1-p} dp = \partial_\tau (y) d\tau + \partial_p (y) dp,$$

and the first-order change in the marginal tax rate at income $y$ is

$$\tilde{T}' (y) - T' (y) = y^{-p} d\tau + (1 - \tau) y^{-p} \ln y dp = \partial_\tau' (y) d\tau + \partial_p' (y) dp.$$

Thus, letting $f_y (y)$ denote the stationary density of incomes, we have

$$\int_0^\infty \left[ \partial_\tau (y) - \frac{T' (y)}{1 - T' (y)} \frac{\varepsilon}{1 + p \varepsilon} y \partial_\tau' (y) \right] f_y (y) \, dy = \frac{1}{1-\tau} \left\{ \frac{1 + \varepsilon}{1 + p \varepsilon} \frac{1}{1-p} \mathbb{E} \left[ y^{1-p} \right] - \frac{\varepsilon}{1 + p \varepsilon} \mathbb{E} \left[ y \ln y \right] \right\},$$

$$\int_0^\infty \left[ \partial_p (y) - \frac{T' (y)}{1 - T' (y)} \frac{\varepsilon}{1 + p \varepsilon} y \partial_p' (y) \right] f_y (y) \, dy = - \frac{1}{1-p} \frac{1-\tau}{1-p} \mathbb{E} \left[ y^{1-p} \right] - \frac{\varepsilon}{1 + p \varepsilon} \mathbb{E} \left[ y \ln y \right] + \frac{1 + \varepsilon}{1 + p \varepsilon} \frac{1-\tau}{1-p} \mathbb{E} \left[ y^{1-p} \ln y \right].$$

In the frictionless model, the income distribution $f_y (y)$ is double-Pareto-lognormal with parameters $(r_{1,\delta}^{\rho_2}, r_{2,\delta}^{\rho_2}, m_y, s_y^2)$, where $m_y = \frac{1+\varepsilon}{1+p\varepsilon} m_w + \frac{\varepsilon \ln (1-\tau)}{1+p\varepsilon}$ and $s_y = \frac{1+\varepsilon}{1+p\varepsilon} s_w$. We can thus derive directly formulas (59) and (60). The parameters of the economy are affected by the perturbation
in the following way:

\[
\frac{dm_y}{d\tau} = -\frac{1}{1 - \tau + pe} \frac{\varepsilon}{1 - \tau + pe}, \quad \frac{d}{d\tau} \left\{ s_y, \mu_y, \sigma_y, r'_{1,y}, r'_{2,y} \right\} = 0, \\
\frac{d}{dp} \left\{ m_y, s_y, \mu_y, \sigma_y, r'_{1,y}, r'_{2,y} \right\} = \frac{\varepsilon}{1 + pe} \left\{ -m_y - s_y - \mu_y - \sigma_y, r'_{1,y}, r'_{2,y} \right\}.
\]

Thus the density of incomes satisfies, letting \( \Phi_i \equiv \Phi \left( \frac{\ln y - m_y}{s_y} + r'_{i,y} s_y \right) \) for \( i \in \{1, 2\} \) with similar definitions of \( \Phi_i^c \) and \( \varphi_1 \), and \( e_i \equiv \frac{|r'_{1,y}|^2}{r'_{1,y} + r'_{2,y}} e^{\frac{1}{2} (r'_{1,y})^2} y^{r'_{1,y} - 1} \Phi_1 + e^{\frac{1}{2} (r'_{2,y})^2} y^{r'_{2,y} - 1} \Phi_2 \).

\[
f_y(y) = \frac{|r'_{1,y}|^2}{r'_{1,y} + r'_{2,y}} \left\{ e^{\frac{1}{2} (r'_{1,y})^2} y^{r'_{1,y} - 1} \Phi_1 + e^{\frac{1}{2} (r'_{2,y})^2} y^{r'_{2,y} - 1} \Phi_2 \right\},
\]

\[
\frac{df_y(y)}{d\tau} = \frac{1}{1 - \tau + pe} \frac{\varepsilon}{1 + pe} \frac{|r'_{1,y}|^2}{r'_{1,y} + r'_{2,y}} \left\{ e^{\frac{1}{2} (r'_{1,y})^2} y^{r'_{1,y} - 1} \Phi_1 + e^{\frac{1}{2} (r'_{2,y})^2} y^{r'_{2,y} - 1} \Phi_2 \right\},
\]

\[
\frac{df_y(y)}{dp} = \frac{\varepsilon}{1 + pe} \frac{|r'_{1,y}|^2}{r'_{1,y} + r'_{2,y}} \left\{ e^{\frac{1}{2} (r'_{1,y})^2} y^{r'_{1,y} - 1} \Phi_1 + e^{\frac{1}{2} (r'_{2,y})^2} y^{r'_{2,y} - 1} \Phi_2 \right\}.
\]

The first-order change in tax revenue due to a perturbation \( d\tau \) in the frictionless model, \( \frac{d\hat{\Phi}^s(T)}{d\tau} \), is given by

\[
\frac{d\hat{\Phi}^s}{d\tau} = \frac{d}{d\tau} \left\{ \int_0^\infty \left( y - \frac{1 - \tau}{1 - p} y^{1-p} \right) f_y(y) dy \right\} = \frac{1}{1 - p} \int_0^\infty y^{1-p} f_y(y) dy + \int_0^\infty \left( y - \frac{1 - \tau}{1 - p} y^{1-p} \right) \frac{df_y(y)}{d\tau} dy
\]

\[
= - \frac{1}{1 - \tau + pe} \mathbb{E}[y] + \frac{1}{1 + \tau} \frac{\varepsilon}{1 + pe} \mathbb{E}[y^{1-p}],
\]

where the last line is obtained by integrating by parts to compute the integrals of the form \( \int_0^\infty e^{(r'_{i,y} - 1 + \alpha)} \ln y \frac{1}{s_y} \varphi_i dy \) with \( r'_{i,y} + \alpha < 0 \). Similarly, the effect of a perturbation \( dp \) is given by

\[
\frac{d\hat{\Phi}^s}{dp} = \int_0^\infty \left( -\frac{1}{1 - p} - \frac{1 - \tau}{1 - p} y^{1-p} + \frac{1 - \tau}{1 - p} y^{1-p} \ln y \right) f_y(y) dy + \int_0^\infty \left( y - \frac{1 - \tau}{1 - p} y^{1-p} \right) \frac{df_y(y)}{dp} dy
\]

\[
= - \frac{1}{1 - p} \frac{1 - \tau}{1 - p} \mathbb{E}[y^{1-p}] - \frac{\varepsilon}{1 + pe} \mathbb{E}[y \ln y] + \frac{1}{1 + pe} \frac{1 - \tau}{1 - p} \mathbb{E}[y^{1-p} \ln y],
\]

again integrating by parts to compute the integrals of the form \( \int_0^\infty e^{(r'_{i,y} - 1 + \alpha)} \ln y \frac{1}{s_y} \varphi_i dy \). Note that a different way of showing these results is to use the KFE characterization of the income
distribution to deduce that
\[
\begin{align*}
  f_y^{\tau + d\tau} (y) & = \lim_{d\tau \to 0} \left( 1 + \frac{\varepsilon}{1 + p\varepsilon \tau} \frac{d\tau}{1 - \tau} \right) f_y^{\tau} \left( 1 + \frac{\varepsilon}{1 + p\varepsilon \tau} \frac{d\tau}{1 - \tau} \right) y + o(d\tau), \\
  f_y^{p + dp} (y) & = \lim_{dp \to 0} \left( 1 + \frac{\varepsilon}{1 + p\varepsilon \tau} dp \right) y^{\tau + dp} f_y^{p} \left( y^{1 + \tau + dp} \right) + o(dp),
\end{align*}
\]
which mean that the individuals with income \( y \) before the perturbation \( d\tau \) (resp., \( dp \)) end up earning \( y' = \left( 1 - \frac{\varepsilon}{1 + p\varepsilon} \frac{d\tau}{1 - \tau} \right) y \) (resp., \( y' = y^{1 + \frac{\varepsilon}{1 + p\varepsilon}} dp \)) after the perturbation, so that \( f_y^{\tau + d\tau} (y) dy' = f_y^{\tau} (y) dy \) (resp., \( f_y^{p + dp} (y') dy' = f_y^{p} (y) dy \)). It is then straightforward (with a change of variables in the integral) to obtain the formulas above:
\[
\begin{align*}
\frac{d\mathcal{K}^*}{d\tau} &= \frac{d}{d\tau} \left\{ \int_0^\infty \left( y - \frac{1 - \tau - d\tau}{1 - p} y^{1-p} \right) f_y^{\tau + d\tau} \left( y^{1 + \tau + dp} \right) y dy - \mathcal{K} \right\} \\
&= \frac{1}{d\tau} \left\{ \int_0^\infty \frac{1}{1-\gamma} \left( \frac{1 + p\varepsilon}{1 + p\varepsilon - \frac{d\tau}{1 - \tau}} y^{1-p} \right)^{1-\gamma} f_y^{\tau + d\tau} (y) dy - \int_0^\infty \mathcal{V} (y) f_y^{\tau} (y) dy \right\} \\
&= \frac{1}{d\tau} \left\{ \int_0^\infty \frac{1}{1-\gamma} \left( \frac{1 + p\varepsilon}{1 + p\varepsilon - \frac{d\tau}{1 - \tau}} \left( 1 - \frac{\varepsilon}{1 + p\varepsilon} \frac{d\tau}{1 - \tau} \right) \right)^{1-\gamma} \right. \\
&\quad \left. \frac{f_y^{\tau} (y) dy}{\rho + \beta - (1 - \gamma) \mu_c - \frac{1}{2} (1 - \gamma)^2 \sigma^2_y} \right\} \\
&= - \int_0^\infty \frac{1}{\rho + \beta - (1 - \gamma) \mu_c - \frac{1}{2} (1 - \gamma)^2 \sigma^2_y} \left( \frac{1 + p\varepsilon}{1 + p\varepsilon - \frac{d\tau}{1 - \tau}} y^{1-p} \right)^{1-\gamma} f_y^{\tau} (y) dy \\
&= - \lambda \int_0^\infty \omega^* (y) \partial_\tau (y) f_y^{\tau} (y) dy,
\end{align*}
\]
and similarly for a perturbation $dp$:

$$
\frac{d\mathcal{W}^* (T)}{dp} = \frac{1}{dp} \left\{ \int_0^\infty \frac{1}{1-\gamma} \left( \frac{1+(p+dp)\varepsilon}{1+\varepsilon} - \frac{1}{1-(p+dp)} y^{1-(p+dp)} \right) \left( \frac{1}{1-p} \frac{1+\varepsilon}{1+p} dp \right) \mu_c - \frac{1}{2} (1-\gamma)^2 \left( \frac{1}{1-p} \frac{1+\varepsilon}{1+p} dp \right)^2 \sigma_e^2 \right\} \times \ldots \times f_y^{p+dp} (y) dy - \mathcal{W}^* (T) \right\}
$$

$$
= \frac{1}{dp} \left\{ \left( 1 + \frac{1-\gamma}{1-p} \frac{1+\varepsilon}{1+\varepsilon} dp - \frac{1}{1-\gamma} \left( \frac{1}{1-p} \frac{1+\varepsilon}{1+p} dp \right) \right) \mu_c - \frac{1}{2} (1-\gamma)^2 \sigma_e^2 dp \right\} \times \ldots
$$

$$
\ldots \times \int_0^\infty \frac{1}{1-\gamma} \left( \frac{1}{1-p} \frac{1+\varepsilon}{1+p} \ln y dp \right) \left( \frac{1}{1-p} \frac{1+\varepsilon}{1+p} \right) f_y (y) dy - \mathcal{W}^* (T)
$$

$$
= - \lambda^* \int_0^\infty \omega^* (y) \partial_p (y) f_y (y) dy - \lambda^* \int_0^\infty \frac{dp}{dp} \omega^* (y) f_y (y) dy.
$$

We finally obtain the optimal tax schedule by imposing that

$$
0 = \frac{d\mathcal{W}^*}{d\tau} + \lambda^* \frac{d\mathcal{R}^*}{d\tau} = \frac{d\mathcal{W}^*}{dp} + \lambda^* \frac{d\mathcal{R}^*}{dp},
$$

which proves formulas (59) and (60) in the frictionless model.

Next I first formulas (51) and (59) in the frictional model.

**Proof of Propositions 5 and 6.** I show that a perturbation $d\tau$ of the tax schedule has the following first-order effects on the density functions in the frictional model: for all $u, \delta$,

$$
f^{x,\tau+d\tau}_{lny,\delta} (u, \delta) = f^{x,\tau}_{lny,\delta} \left( u + \frac{\varepsilon}{1+\varepsilon} \frac{d\tau}{1-\tau}, \delta \right)
$$

i.e.,

$$
f^{x,\tau+d\tau}_{y,\delta} \left( y, \delta \right) = \partial_{d\tau} \left( 1 + \frac{\varepsilon}{1+\varepsilon} \frac{d\tau}{1-\tau} \right) f^{x,\tau}_{y,\delta} \left( 1 + \frac{\varepsilon}{1+\varepsilon} \frac{d\tau}{1-\tau} \right),
$$

for $x \in \{i, s\}$. To see this, consider the functions

$$
g^x (u, \delta) \equiv f^{x,\tau}_{lny,\delta} \left( u + \frac{\varepsilon}{1+\varepsilon} \frac{d\tau}{1-\tau}, \delta \right).
$$

I show that $g^i, g^s$ satisfy the KFE and boundary conditions that define the functions $f^{i,\tau+d\tau}_{lny,\delta}$ and $f^{s,\tau+d\tau}_{y,\delta}$, respectively, which will imply the result. First, note that $\mu_y, \sigma_y, \mu_\delta, \sigma_\delta, \delta^*, \delta$ do not depend on $\tau$. We have, for all $u \in \mathbb{R}$, all $\delta \in (\tilde{\delta}, \delta^*) \cup (\delta^*, \tilde{\delta})$ if $x = i$, and all $\delta \in \mathbb{R} \setminus \{\tilde{\delta}, \delta\}$ if $x = s$, letting
\( \hat{u} \equiv u + \frac{\varepsilon}{1 - \varepsilon \tau} \frac{d\tau}{1 - \tau}, \)

\[-(\beta + q f^s) g(u, \delta) - \mu_y g_1(u, \delta) + \mu_\delta g_2(u, \delta) + \frac{1}{2} \sigma'_y \gamma g_1(u, \delta) + \frac{1}{2} \sigma'_\delta g_22(u, \delta) - \sigma_y \sigma_\delta g_12(u, \delta) = - (\beta + q \hat{f}^\tau ) f^{x\tau}_{\ln y^\tau, \delta}(\hat{u}, \delta) - \mu_y \partial_1 f^{x\tau}_{\ln y^\tau, \delta}(\hat{u}, \delta) + \mu_\delta \partial_2 f^{x\tau}_{\ln y^\tau, \delta}(\hat{u}, \delta) + \frac{1}{2} \sigma'_y \partial_1 \gamma_{\ln y^\tau, \delta}(\hat{u}, \delta) + \frac{1}{2} \sigma'_\delta \partial_2 \gamma_{\ln y^\tau, \delta}(\hat{u}, \delta) - \sigma_y \sigma_\delta \partial_12 \gamma_{\ln y^\tau, \delta}(\hat{u}, \delta) = 0, \]

where the last equality follows from the KFE solved by \( f^{x\tau}_{\ln y^\tau, \delta} \), evaluated at \((\hat{u}, \delta)\). Next, note that the density of incomes at birth satisfies

\[ f^{\tau+dr}_{\ln y^\tau_0}(u) = \frac{1}{s_y \sqrt{2\pi}} e^{-\frac{1}{2s_y^2} \left( u - \frac{\varepsilon}{1 - \varepsilon \tau} \frac{d\tau}{1 - \tau} \right)^2} = f^{\tau}_{\ln y^\tau_0}(u + \frac{\varepsilon}{1 + \varepsilon \tau} \frac{d\tau}{1 - \tau}). \]

Thus we have

\[ g_1^s(u, \delta^s) - g_1^s(u, \delta^-) - \frac{1 - p}{1 + 1/\varepsilon} (g_2^s(u, \delta^s) - g_2^s(u, \delta^-)) - \left( g_1^s(u, \delta^s) - \frac{1 - p}{1 + 1/\varepsilon} g_2^s(u, \delta^-) \right) = \partial_1 f^{s,\tau}_{\ln y^\tau, \delta}(\hat{u}, \delta^+) - \partial_1 f^{s,\tau}_{\ln y^\tau, \delta}(\hat{u}, \delta^-) - \frac{1 - p}{1 + 1/\varepsilon} \left( \partial_2 f^{s,\tau}_{\ln y^\tau, \delta}(\hat{u}, \delta^+) - \partial_2 f^{s,\tau}_{\ln y^\tau, \delta}(\hat{u}, \delta^-) \right) \]

\[ = \partial_1 f^{s,\tau}_{\ln y^\tau, \delta}(\hat{u}, \delta^-) - \frac{1 - p}{1 + 1/\varepsilon} \partial_2 f^{s,\tau}_{\ln y^\tau, \delta}(\hat{u}, \delta^-) = 0, \]

where the last equality follows from equation (45) satisfied by \( f^{\tau}_{\ln y^\tau, \delta} \), evaluated at \((\hat{u}, \delta)\). The corresponding equation (44) for \( g^i, g^s \) is shown in the same way. Similarly,

\[ g_1^i(u, \delta^s) - g_1^i(u, \delta^-) - \frac{1 - p}{1 + 1/\varepsilon} (g_2^i(u, \delta^s) - g_2^i(u, \delta^-)) - \frac{2}{\sigma_y \sigma_\delta} \left( \beta f^{\tau}_{\ln y^\tau_0}(\hat{u}) + q f^{s,\tau}(\hat{u}) \right) = \partial_1 f^{i,\tau}_{\ln y^\tau, \delta}(\hat{u}, \delta^s) - \partial_1 f^{i,\tau}_{\ln y^\tau, \delta}(\hat{u}, \delta^-) - \frac{1 - p}{1 + 1/\varepsilon} \left( \partial_2 f^{i,\tau}_{\ln y^\tau, \delta}(\hat{u}, \delta^s) - \partial_2 f^{i,\tau}_{\ln y^\tau, \delta}(\hat{u}, \delta^-) \right) \]

\[ = \frac{2}{\sigma_y \sigma_\delta} \left( \beta f^{\tau}_{\ln y^\tau_0}(\hat{u}) + q f^{s,\tau}(\hat{u}) \right) = 0, \]

where the last equality follows from the third conservation law satisfied by \( f^{\tau}_{\ln y^\tau, \delta} \), evaluated at \((\hat{u}, \delta)\). Finally check the other boundary conditions: we have \( g^i(u, \delta) = f^{i,\tau}_{\ln y^\tau, \delta}(\hat{u}, \delta) = 0 \) and similarly \( g^i(u, \delta) = 0 \), where the last equalities follow from the corresponding boundary conditions of \( f^i \). Similarly, we have, for all \( h \in \mathbb{R} \),

\[ \lim_{\delta \to \pm \infty} g^s \left( h - \frac{1 + 1/\varepsilon}{1 - p} \delta, \delta \right) = f^{s,\tau}_{\ln y^\tau, \delta} \left( h + \frac{\varepsilon}{1 + \varepsilon \tau} \frac{d\tau}{1 - \tau} - \frac{1 + 1/\varepsilon}{1 - p} \delta, \delta \right) = 0. \]

Finally, we have \( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \{ g^i + g^s \} (u, \delta) dud\delta = 1 \), which completes the proof that \( g^i = f^{i,\tau+dr}_{\ln y^\tau, \delta} \) and \( g^s = f^{s,\tau+dr}_{\ln y^\tau, \delta} \).
Equation (64) implies that, for $x \in \{i, s\}$, for all $v, \delta$,

$$f^{x, \tau + d\tau}_y (v, \delta) = f^{x, \tau + d\tau}_y (v - \delta, \delta) = f^{x, \tau}_y \left( v + \frac{x}{1 + p \varepsilon} \frac{d\tau}{1 - \tau} - \delta, \delta \right) = f^{x, \tau}_y \left( v + \frac{x}{1 + p \varepsilon} \frac{d\tau}{1 - \tau}, \delta \right),$$

and thus the following relationship between the marginal densities of income given taxes $\tau$ and $\tau + d\tau$ holds:

$$f^{\tau + d\tau}_y (v) = \int_{-\infty}^{\infty} \left\{ f^{i, \tau + d\tau}_y (v, \delta) + f^{s, \tau + d\tau}_y (v, \delta) \right\} d\delta$$

$$= \int_{-\infty}^{\infty} \left\{ f^{i, \tau}_y \left( v + \varepsilon \frac{d\tau}{1 + p \varepsilon} - \delta, \delta \right) + f^{s, \tau}_y \left( v + \varepsilon \frac{d\tau}{1 + p \varepsilon} - \delta, \delta \right) \right\} d\delta = f^{\tau}_y \left( v + \varepsilon \frac{d\tau}{1 + p \varepsilon} - \tau, \delta \right).$$

We therefore find, with the same change of variables as in the previous proof:

$$\frac{d\mathcal{R}}{d\tau} = \frac{d}{d\tau} \left\{ \int_{-\infty}^{\infty} \left( e^u - \frac{1 - \tau}{1 - p} e^{(1-p)u} \right) f_{y \delta} (u) du \right\}$$

$$= \frac{1}{d\tau} \left\{ \int_{-\infty}^{\infty} \left( e^u - \frac{1 - \tau - d\tau}{1 - p} e^{(1-p)u} \right) f_{y \delta} (u + \varepsilon \frac{d\tau}{1 + p \varepsilon} - \delta) du - \mathcal{R} \right\}$$

$$= \int_{-\infty}^{\infty} \left( -\frac{\varepsilon}{1 + p \varepsilon - \tau} e^u + \frac{1 + \varepsilon}{1 + p \varepsilon - 1 - \tau} e^{(1-p)u} \right) f_{y \delta} (u) du$$

$$= -\frac{\varepsilon}{1 + p \varepsilon - \tau} \mathbb{E} [y] + \frac{1}{1 + p \varepsilon - 1 - \tau} \mathbb{E} [y^{1-p}].$$

Note that the same computations (keeping only the term $e^u$ in the integral) imply that $\frac{d}{d\tau} \mathbb{E} [y] = -\frac{\varepsilon}{1 + p \varepsilon - 1 - \tau} \mathbb{E} [y]$, proving equation (51). Finally,

$$\frac{d\mathcal{W}}{d\tau} = \frac{d}{d\tau} \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \nu_i (e^u, \delta) f_{y \delta}^i (u, \delta) dud\delta + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \nu_s (e^u, \delta) f_{y \delta}^s (u, \delta) dud\delta \right\}$$

$$= \sum_{x \in \{i, s\}} \frac{1}{d\tau} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left\{ \nu_i (e^u, \delta) \frac{d\tau}{1 + p \varepsilon - 1 - \tau} \right\} \tilde{\nu}_x (\delta) f_{y \delta}^{x, (\tau)} (v, \delta) dv \left( \frac{d\tau}{1 + p \varepsilon - 1 - \tau} \right)$$

$$= -\lambda \int_{0}^{\infty} \omega (y) \partial_y (y) f_{y} (y) dy.$$

This concludes the proof of Proposition 6.

I finally prove equation (61).

**Proof of Proposition 7.** Suppose first that the following assumption (*) holds:

$$\{\delta, \delta^*, \delta\} (p + dp) = \left( 1 - \frac{1 + \varepsilon}{1 - p \varepsilon} \right) \{\delta, \delta^*, \delta\} = \frac{\varepsilon \sigma_p (p + dp)}{\sigma_\delta} \{\delta, \delta^*, \delta\}.$$
Then the perturbation $dp$ affects the densities of income as follows:

\[
\begin{align*}
 f_{\ln y^*, \delta}^{x, p + dp} (u, \delta) &= \left(1 + \frac{\varepsilon}{1 + p \varepsilon} dp\right) \left(1 + \frac{1}{1 - p} \frac{1 + \varepsilon}{1 + p \varepsilon} dp\right) \times \ldots \\
 f_{\ln y^*, \delta}^{x, p} \left(\left(1 + \frac{\varepsilon}{1 + p \varepsilon} dp\right) u, \left(1 + \frac{1}{1 - p} \frac{1 + \varepsilon}{1 + p \varepsilon} dp\right) \delta\right) &\equiv g^x (u, \delta),
\end{align*}
\]

(65)

for $x \in \{i, s\}$. To see this, I show that $g^i, g^s$ satisfy the KFE and boundary conditions that define the functions $f_{\ln y^*, \delta}^{i, p + dp}$ and $f_{\ln y^*, \delta}^{s, p + dp}$, respectively.

For any $\delta, u$, let $\bar{\delta} \equiv \left(1 - \frac{1}{1 - p} \frac{1 + \varepsilon}{1 + p \varepsilon} dp\right) \delta$, $\bar{\delta} \equiv \left(1 - \frac{1}{1 - p} \frac{1 + \varepsilon}{1 + p \varepsilon} dp\right) \delta$, and $\bar{u} = \left(1 + \frac{\varepsilon}{1 + p \varepsilon} dp\right) u$. First, we have, for all $u \in \mathbb{R}$, all $\delta \in (\bar{\delta}, \delta^*) \cup (\bar{\delta}^*, \bar{\delta})$ if $x = i$, and all $\delta \in \mathbb{R} \setminus \{\bar{\delta}, \bar{\delta}\}$ if $x = s$,

\[
- (\beta + q f^i) g^x - \left(1 - \frac{\varepsilon}{1 + p \varepsilon} dp\right) \mu_y g^i_1 + \frac{1 - p}{1 + \varepsilon} \left(1 - \frac{1}{1 - p} \frac{1 + \varepsilon}{1 + p \varepsilon} dp\right) \left(1 - \frac{\varepsilon}{1 + p \varepsilon} dp\right) \mu_y g^x_2
\]

\[
+ \frac{1}{2} \left(1 - \frac{\varepsilon}{1 + p \varepsilon} dp\right)^2 \sigma^2_y g^i_1 + \frac{1}{2} \left(1 - \frac{1}{1 + \varepsilon} dp\right)^2 \left(1 - \frac{1}{1 - p} \frac{1 + \varepsilon}{1 + p \varepsilon} dp\right)^2 \sigma^2_y g^x_2
\]

\[
- \frac{1 - p}{1 + \varepsilon} \left(1 - \frac{1}{1 - p} \frac{1 + \varepsilon}{1 + p \varepsilon} dp\right)^2 \sigma^2_y g^i_2
\]

\[
= \left(1 + \frac{\varepsilon}{1 + p \varepsilon} dp\right) \left(1 + \frac{1}{1 - p} \frac{1 + \varepsilon}{1 + p \varepsilon} dp\right) \left\{- (\beta + q f^i) f_{\ln y^*, \delta}^{i, (p)} - \mu_y \frac{\partial f_{\ln y^*, \delta}^{i, p}}{\partial u} + \frac{1 - p}{1 + \varepsilon} \frac{\partial f_{\ln y^*, \delta}^{i, p}}{\partial \delta} \right\}_{(\bar{u}, \bar{\delta})} = 0,
\]

where the last equality follows from the KFE solved by $f_{\ln y^*, \delta}^{i, p}$, evaluated at $(\bar{u}, \bar{\delta})$. Thus $g^x$ satisfies the KFE of $f_{\ln y^*, \delta}^{i, p}$ for all $\delta \in (\bar{\delta}, \delta^*) \cup (\bar{\delta}^*, \bar{\delta})$ if $x = i$ and $\bar{\delta} \in \mathbb{R} \setminus \{\bar{\delta}, \bar{\delta}\}$ if $x = s$, i.e. the same domains as those of $f_{\ln y^*, \delta}^{i, p}$. Next, we have, for $\delta \in \{\bar{\delta}, \bar{\delta}\}$,

\[
g^i (u, \delta^+) - g^i (u, \delta^-) - \frac{1 - p}{1 + 1/\varepsilon} \left(1 - \frac{dp}{1 - p}\right) \left(\hat{g}_2^i (u, \delta^+) - \hat{g}_2^i (u, \delta^-)\right)
\]

\[
- \left(\hat{g}_1^i (u, \delta^-) - \frac{1 - p}{1 + 1/\varepsilon} \left(1 - \frac{dp}{1 - p}\right) \hat{g}_2^i (u, \delta^-)\right)
\]

\[
= \left(1 + \frac{\varepsilon}{1 + p \varepsilon} dp\right)^2 \left(1 + \frac{1}{1 - p} \frac{1 + \varepsilon}{1 + p \varepsilon} dp\right) \left\{\frac{\partial f_{\ln y^*, \delta}^{i, p}}{\partial u} \left(\bar{u}, \bar{\delta}^+\right) - \frac{\partial f_{\ln y^*, \delta}^{i, p}}{\partial u} \left(\bar{u}, \bar{\delta}^-\right)\right\}
\]

\[
- \frac{1 - p}{1 + 1/\varepsilon} \left(1 - \frac{dp}{1 - p}\right) \left(1 + \frac{\varepsilon}{1 + p \varepsilon} dp\right) \left(1 + \frac{1}{1 - p} \frac{1 + \varepsilon}{1 + p \varepsilon} dp\right)^2 \left\{\frac{\partial f_{\ln y^*, \delta}^{i, p}}{\partial \delta} \left(\bar{u}, \bar{\delta}^+\right) - \frac{\partial f_{\ln y^*, \delta}^{i, p}}{\partial \delta} \left(\bar{u}, \bar{\delta}^-\right)\right\}
\]

\[
- \left(1 + \frac{\varepsilon}{1 + p \varepsilon} dp\right)^2 \left(1 + \frac{1}{1 - p} \frac{1 + \varepsilon}{1 + p \varepsilon} dp\right) \left\{\frac{\partial f_{\ln y^*, \delta}^{i, p}}{\partial u} \left(\bar{u}, \bar{\delta}^-\right) - \frac{1 - p}{1 + 1/\varepsilon} \frac{\partial f_{\ln y^*, \delta}^{i, p}}{\partial \delta} \left(\bar{u}, \bar{\delta}^-\right)\right\} = 0,
\]

where the last equality follows from the corresponding conservation law satisfied by $f_{\ln y^*, \delta}^{i, p}$, evaluated
at \((\hat{u}, \hat{\delta})\), and noting that

\[
\beta f^{p+dp}_{ln y_{0}}(u) + q f^{s,p+dp}_{ln y^*}(u) = \frac{\beta}{(1 - \frac{1}{1+pe} dp)} \frac{e^{-\frac{1}{1+pe} dp}}{s y \sqrt{2\pi}} (u - (1 - \frac{1}{1+pe} dp) m_y)^2
\]

\[
+ q \left(1 + \frac{\varepsilon}{1+pe} dp\right) \left(1 + \frac{1}{1-p} + \varepsilon dp\right) \int_{(p+dp)}^{\delta} f^{s,p}_{ln y^*,\delta} (\hat{u}, \hat{\delta}) d\delta,
\]

we obtain

\[
g_{i}^{(u, \delta^+)} - g_{i}^{(u, \delta^-)} - \frac{1-p}{1+1/\varepsilon} \left(1 - \frac{dp}{1-p}\right) \left(g_{2}^{(u, \delta^+)} - g_{2}^{(u, \delta^-)}\right)
\]

\[
- \frac{\beta}{s y \sqrt{2\pi}} \frac{1}{s y \sqrt{2\pi}} (\hat{u} - m_y)^2 q \int_{\delta}^{\hat{\delta}} f^{s,p}_{ln y^*,\delta} (\hat{u}, \hat{\delta}) d\delta
\]

\[
= \left(1 + \frac{\varepsilon}{1+pe} dp\right) \left(1 + \frac{1}{1-p} + \varepsilon dp\right) \left(1 - \frac{1-p}{1+1/\varepsilon} \left(1 + \frac{1+1/\varepsilon}{1+pe} dp\right) \left(\frac{\partial f^{i,p}_{ln y^*,\delta}}{\partial u} (\hat{u}, \hat{\delta}^+) - \frac{\partial f^{i,p}_{ln y^*,\delta}}{\partial u} (\hat{u}, \hat{\delta}^-)\right)\right)
\]

\[
- \frac{1-p}{1+1/\varepsilon} \left(1 + \frac{\varepsilon}{1+pe} dp\right) \left(1 - \frac{1+1/\varepsilon}{1+pe} dp\right) \left(\frac{\partial f^{i,p}_{ln y^*,\delta}}{\partial \delta} (\hat{u}, \hat{\delta}^+) - \frac{\partial f^{i,p}_{ln y^*,\delta}}{\partial \delta} (\hat{u}, \hat{\delta}^-)\right)
\]

\[
- \frac{1}{(1 - \frac{dp}{1-p}) \left(1 - \frac{1}{1+pe} dp\right)} \frac{2}{\sigma y \sigma \delta} \frac{\beta e^{-\frac{1}{2s y} (\hat{u} - m_y)^2}}{s y \sqrt{2\pi}} q f^{s,p}_{ln y^*} (\hat{u}) = 0,
\]

where the last equality follows from the corresponding conservation law satisfied by \(f^{p}_{ln y^*,\delta}\), evaluated at \((\hat{u}, \hat{\delta})\). The remaining boundary conditions are proved similarly, which concludes the proof.

Note that equation (65) implies

\[
f^{(p+dp)}_{y,\delta} (y, \delta) = e^{-\frac{\varepsilon}{1+pe} dp} f^{(p+dp)}_{y^*,\delta} (y e^{-\delta}, \delta) = \left(1 + \frac{\varepsilon}{1+pe} dp\right) \left(1 + \frac{1+1/\varepsilon}{1+pe} dp\right) e^{\frac{\varepsilon}{1+pe} ln y_{dp}} \times \ldots
\]

\[
\times \ldots e^{\frac{1}{1-p} \frac{1+1/\varepsilon}{1+pe} - \frac{\varepsilon}{1+pe} dp} \delta dp f^{(p)}_{y,\delta} \left(e^{\frac{1+1/\varepsilon}{1+pe} dp} \ln y + \left[\frac{1+1/\varepsilon}{1+pe} - \frac{\varepsilon}{1+pe} \right] \delta dp, \left(1 + \frac{1}{1-p} + \varepsilon dp\right) \delta\right).
\]

Second, I compute the effect of a perturbation \(dp\) on government revenue:

\[
\frac{d\mathcal{R}}{dp} = \frac{d}{dp} \left\{ \int_{0}^{\infty} \left( y - \frac{1}{1-p} y^{1-p} \right) f^{p}_{y} (y) dy \right\}
\]

\[
= \int_{0}^{\infty} \left( \ln y - \frac{1}{1-p} \right) \frac{1}{1-p} y^{1-p} f^{p}_{y} (y) dy + \frac{1}{dp} \int_{0}^{\infty} \left( y - \frac{1}{1-p} y^{1-p} \right) \left\{ f^{p+dp}_{y} (y) - f^{p}_{y} (y) \right\} dy.
\]

The first term in the right hand side is the standard mechanical effect \(M\) of the perturbation, as in the frictionless model. I now decompose the second integral into three parts: \(\frac{d\mathcal{R}}{dp} = M + \)
\[
\frac{1}{\delta p} (B_1 + B_2 + B_3), \text{ where}
\]
\[
B_1 = \int_0^\infty \int_{-\infty}^\infty T(y) \left\{ \left( 1 + \frac{\varepsilon}{1 + \delta p} \right)^2 y^{\frac{\varepsilon}{1 + \delta p}} f_p p_{y,\delta} \left( y^{1 + \frac{\varepsilon}{1 + \delta p}}, 1 + \frac{\varepsilon}{1 + \delta p} \right) \delta - \ldots 
\right. 
\]
\[
\left. \ldots - f_p p_{y,\delta} (y, \delta) \right\} dyd\delta
\]
\[
B_2 = \int_0^\infty \int_{-\infty}^\infty T(y) \left\{ \left( 1 + \frac{\varepsilon}{1 + \delta p} \right) \left( 1 + \frac{1 + \varepsilon}{1 - p 1 + \delta p} \right) e^{\frac{\varepsilon}{1 + \delta p} \ln y \delta p} e^{\left[ \frac{1 + \varepsilon}{1 + \delta p} - \frac{\varepsilon}{1 + \delta p} \right] \delta p} \times \ldots 
\right. 
\]
\[
\left. \ldots \times f_p p_{y,\delta} \left( y^{1 + \frac{\varepsilon}{1 + \delta p}}, 1 + \frac{\varepsilon}{1 + \delta p} \right) \delta \right\} dyd\delta
\]
\[
B_3 = \int_0^\infty \int_{-\infty}^\infty T(y) \left\{ f_p p_{y,\delta} (y, \delta) - \left( 1 + \frac{\varepsilon}{1 + \delta p} \right) \left( 1 + \frac{1 + \varepsilon}{1 - p 1 + \delta p} \right) e^{\frac{\varepsilon}{1 + \delta p} \ln y \delta p} \times \ldots 
\right. 
\]
\[
\left. \times e^{\left[ \frac{1 + \varepsilon}{1 - p 1 + \delta p} - \frac{\varepsilon}{1 + \delta p} \right] \delta p} f_p p_{y,\delta} \left( y^{1 + \frac{\varepsilon}{1 + \delta p}}, 1 + \frac{\varepsilon}{1 + \delta p} \right) \delta \right\} dyd\delta
\]

and I compute each term in turn.

\([B_1]\) : We have
\[
B_1 = \int_0^\infty \left( \left( 1 - \frac{\varepsilon}{1 + \delta p} \right) y - \frac{1 - \tau}{1 - p} \left( 1 - \frac{\varepsilon (1 - p)}{1 + \delta p} \ln y \delta p \right) y^{1 - p} \right) f_p p_{y} (y) dy
\]
\[
- \int_0^\infty \left( y - \frac{1 - \tau}{1 - p} y^{1 - p} \right) f_p p_{y} (y) dy
\]
\[
= - \frac{\varepsilon}{1 + \delta p} \int_0^\infty (y - (1 - \tau) y^{1 - p}) \ln y f_p p_{y} (y) dy
\]
\[
= - \frac{\varepsilon}{1 + \delta p} \int_0^\infty T'(y) \left[ \frac{y^{1 - p}}{1 - T'(y)} \right] y f_p p_{y} (y) dy,
\]
which is the standard behavioral effect found in frictionless models.

\([B_2]\) : We have
\[
B_2 = \int_0^\infty \int_{-\infty}^\infty \left\{ T \left( y^{1 - \frac{\varepsilon}{1 + \delta p}} e^{\left[ \frac{1 + \varepsilon}{1 - p 1 + \delta p} - \frac{\varepsilon}{1 + \delta p} \right] \delta p} \right) - T \left( y^{1 - \frac{\varepsilon}{1 + \delta p}} \right) \right\} f_p p_{y,\delta} (y, \delta) dyd\delta
\]
\[
= - \left( \frac{1 + \varepsilon}{1 - p 1 + \delta p} - \frac{\varepsilon}{1 + \delta p} \right) \int_0^\infty T'(y) \left[ \int_{-\infty}^\infty \delta f_{y|y} (\delta | y) d\delta \right] y f_p p_{y} (y) dy.
\]
We can show that this term is related to the elasticity of income with respect to a proportional change in the parameters \( \mathbb{Y} = \{ \mu_y, \sigma_y, m_y, s_y \} \) (which all have elasticity \(- \frac{\varepsilon}{1 + \delta p} \) with respect to \( p \)), keeping the parameters \( \mathbb{D} = \{ \mu_\delta, \sigma_\delta, \delta_*, \delta \} \) constant, i.e.,
\[
\left. \frac{\partial \ln (1 - F_p(y))}{\partial \ln y} \right|_{\mathbb{D}} = \frac{\partial \ln y}{\partial \ln \mathbb{Y}}.
\]
We have

\[
B_3 = \int_0^\infty \int_{-\infty}^\infty T(y) \left\{ f_{y,\delta}^{\tilde{\delta},*}(y, \delta) - f_{y,\delta}^p(y, \delta) \right\} dy \delta
\]

\[
= \int_0^\infty \int_{-\infty}^\infty T(y) \left\{ \left( \tilde{\delta} - \delta \right) \frac{\partial f_{y,\delta}(y, \delta)}{\partial \tilde{\delta}} + \left( \tilde{\delta}^* - \delta^* \right) \frac{\partial f_{y,\delta}(y, \delta)}{\partial \delta^*} + \left( \tilde{\delta} - \delta \right) \frac{\partial f_{y,\delta}(y, \delta)}{\partial \delta} \right\} dy \delta
\]

\[
= dp \int_0^\infty \int_{-\infty}^\infty T(y) \left\{ \sum_{\delta_i \in \{\tilde{\delta}, \delta^*, \tilde{\delta}^*, \delta \}} \delta_i \left( \frac{\partial \ln |\delta_i|}{\partial p} - \left( -\frac{1}{1-p} \frac{1 + \varepsilon}{1 + \varepsilon dp} \right) \frac{\partial f_{y,\delta}(y, \delta)}{\partial \delta_i} \right) \right\} dy \delta,
\]

where the first equality follows from the observation that the KFE and boundary conditions defining the density \( f_{x,p+dp}^{\ln y,\delta} \) is solved by the function

\[
f_{x,p+dp}^{\ln y,\delta}(u, \delta) = \left( 1 + \varepsilon \frac{1}{1 + \varepsilon dp} \right) \left( 1 + \frac{1 + \varepsilon}{1 - p} \frac{1 + \varepsilon}{1 + \varepsilon dp} \right) f_{x,\delta}^{\tilde{\delta},*}(\tilde{u}, \delta),
\]

where the r.h.s. is the density given the progressivity \( p \) and the parameters

\[
\left\{ \tilde{\delta}, \delta^*, \tilde{\delta}^*, \delta \right\} \equiv \left( 1 + \frac{1 + p}{1 - p} \frac{1 + \varepsilon}{1 + \varepsilon dp} \right) \left\{ \tilde{\delta} (p + dp), \delta^* (p + dp), \tilde{\delta} (p + dp) \right\},
\]

which implies that

\[
f_{y,\delta}^{x,p+dp}(y, \delta) = \left( 1 + \varepsilon \frac{1}{1 + \varepsilon dp} \right) \left( 1 + \frac{1 + \varepsilon}{1 - p} \frac{1 + \varepsilon}{1 + \varepsilon dp} \right) e^{\frac{y}{1 + \varepsilon dp}} \ln y dp \times \ldots
\]

\[
\ldots \times e^{\frac{1}{1-p} \frac{1 + \varepsilon}{1 + \varepsilon dp} \frac{y}{1 + \varepsilon dp} \delta dp} f_{y,\delta}^{x,\tilde{\delta},*}(y, \delta) \left( e^{\frac{y}{1 + \varepsilon dp} \ln y} e^{\frac{1 + \varepsilon}{1 - p} \frac{1 + \varepsilon}{1 + \varepsilon dp} \delta dp} \right) \left( 1 + \frac{1 + p}{1 - p} \frac{1 + \varepsilon}{1 + \varepsilon dp} \right)
\]

and leads to the expression above after a change of variables. Gathering all the terms, we obtain

\[
\frac{d\mathcal{R}}{dp} = \int_0^\infty \left( \ln y - \frac{1}{1-p} \right) \frac{1 - \tau}{1-p} y^{1-p} f_y(y) dy
\]

\[
- \left( \frac{1}{1-p} \frac{1 + \varepsilon}{1 + \varepsilon dp} - \frac{\varepsilon}{1 + \varepsilon dp} \right) \int_0^\infty T'(y) y E[\delta | y] f_y(y) dy
\]

\[
- \frac{\varepsilon}{1 + \varepsilon dp} \int_0^\infty T'(y) \frac{y \partial_y f_y(y)}{1 - T'(y)} f_y(y) dy
\]

\[
+ \int_0^\infty T(y) \left[ \sum_{i=1}^3 \left( \frac{d \ln |\delta_i|}{dp} - \frac{d \ln |\sigma_i|}{dp} \right) \right] f_y(y) dy
\]

63
Third, I compute the effect of a perturbation $dp$ on the government objective:

$$\frac{d\bar{W}}{dp} = \sum_{x \in \{i, s\}} \frac{d}{dp} \left\{ \int_{0}^{\infty} \int_{-\infty}^{\infty} V_{\pi}^{*}(y) \bar{v}_{x}^{*}(\delta) f_{y,\delta}^{x,p}(y, \delta) \, dyd\delta \right\}$$

$$= \frac{d}{dp} \left\{ \int_{0}^{\infty} \int_{-\infty}^{\infty} \frac{1}{1-\gamma} \left( \frac{1+p+e}{1+p} y^{1-p} \right)^{1-\gamma} \bar{v}^p(y, \delta) f_{y,\delta}^{p}(y, \delta) \, dyd\delta \right\}.$$

I now decompose the integral into three parts: $\frac{d\bar{W}}{dp} = \frac{1}{dp} (W_1 + W_2 + W_3 + W_4)$, where

$$W_1 = \int_{0}^{\infty} \int_{-\infty}^{\infty} \left[ \frac{1}{1-\gamma} \left( \frac{1+p+e}{1+p} y^{1-p} \right)^{1-\gamma} \right. \rho_1 + \rho_2 - (1-\gamma) \left( 1 - \frac{1}{1+p} \frac{1+p+e}{1+p} dp \right) \mu_c - \frac{1}{2} (1-\gamma)^2 \sigma_c^2 \right. \left. \bar{v}^p(y, \delta) f_{y,\delta}^{p}(y, \delta) \right. \, dyd\delta$$

$$W_2 = \int_{0}^{\infty} \int_{-\infty}^{\infty} \left[ \frac{1}{1-\gamma} \left( \frac{1+p+e}{1+p} y^{1-p} \right)^{1-\gamma} \right. \rho_1 + \rho_2 - (1-\gamma) \mu_c - \frac{1}{2} (1-\gamma)^2 \sigma_c^2 \right. \left. \bar{v}^p(y, \delta) f_{y,\delta}^{p}(y, \delta) \right. \, dyd\delta$$

$$W_3 = \int_{0}^{\infty} \int_{-\infty}^{\infty} \left[ \frac{1}{1-\gamma} \left( \frac{1+p+e}{1+p} y^{1-p} \right)^{1-\gamma} \right. \rho_1 + \rho_2 - (1-\gamma) \mu_c - \frac{1}{2} (1-\gamma)^2 \sigma_c^2 \right. \left. \bar{v}^p(y, \delta) f_{y,\delta}^{p}(y, \delta) \right. \, dyd\delta$$

$$W_4 = \int_{0}^{\infty} \int_{-\infty}^{\infty} \left[ \frac{1}{1-\gamma} \left( \frac{1+p+e}{1+p} y^{1-p} \right)^{1-\gamma} \right. \rho_1 + \rho_2 - (1-\gamma) \mu_c - \frac{1}{2} (1-\gamma)^2 \sigma_c^2 \right. \left. \bar{v}^p(y, \delta) f_{y,\delta}^{p}(y, \delta) \right. \, dyd\delta$$

and I compute each term in turn. First, we easily find

$$W_1 + W_2 = \lambda \int_{0}^{\infty} \left[ \partial_{p}(y) \omega(y) + \frac{d \ln \rho}{dp} \bar{v}(y) \right] f_{y}^{p}(y) \, dy,$$

which are the standard (static and dynamic) terms already present in the frictionless formula. Next,
we have

\[ W_3 = \int_0^{\infty} \int_{-\infty}^{\infty} \frac{1}{1 - \nu} \left( \frac{1}{1 + \nu} \frac{1 - \gamma}{1 - \nu} y^{1-\gamma} \right) \times \ldots \]

\[ \times \frac{1}{1 - \nu} \left( \frac{1}{1 + \nu} \frac{1 - \gamma}{1 - \nu} y^{1-\gamma} \right) \times \ldots \]

\[ \ldots \times \left\{ e^{-(1-\gamma)\delta \nu \rho} \left( y c \frac{1}{1 - \nu} \frac{1 - \gamma}{1 - \nu} \delta \nu \rho - \bar{v}(y, \delta) \right) \right\} f_{y,\delta}(y, \delta) \ dyd\delta \]

\[ = - \frac{dp}{1 - p} \int_{-\infty}^{\infty} y \partial_{\nu} \left\{ \frac{1}{1 - \nu} \left( \frac{1}{1 + \nu} \frac{1 - \gamma}{1 - \nu} y^{1-\gamma} \right) \right\} f_{y,\delta}(y, \delta) \ dyd\delta \]

\[ = - \left[ \frac{1}{1 - p} \frac{1 + \varepsilon}{1 + p\varepsilon} \right] dp \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} \delta \nu \partial_{\nu} \bar{v}(y, \delta) \ dy \right\} f_{y,\delta}(y, \delta) \ dy. \]

Finally, we have

\[ W_4 = \int_0^{\infty} \int_{-\infty}^{\infty} \frac{1}{1 - \nu} \left( \frac{1}{1 + \nu} \frac{1 - \gamma}{1 - \nu} y^{1-\gamma} \right) \times \ldots \]

\[ \ldots \times \left( 1 + \frac{\varepsilon}{1 + p\varepsilon} dp \right) e^{\frac{\nu}{1 - \nu} \ln y dp} e^{\left( \frac{1}{1 + \nu} \frac{1 + \gamma}{1 + \nu} \right) \delta \nu \rho} \left\{ f_{y,\delta}(y, \delta) \right\} f_{y,\delta}(y, \delta) \ dyd\delta \]

\[ + \int_0^{\infty} \int_{-\infty}^{\infty} \frac{1}{1 - \nu} \left( \frac{1}{1 + \nu} \frac{1 - \gamma}{1 - \nu} y^{1-\gamma} \right) \times \ldots \]

\[ \ldots \times \left( 1 + \frac{\varepsilon}{1 + p\varepsilon} dp \right) e^{\frac{\nu}{1 - \nu} \ln y dp} e^{\left( \frac{1}{1 + \nu} \frac{1 + \gamma}{1 + \nu} \right) \delta \nu \rho} \left\{ f_{y,\delta}(y, \delta) \right\} f_{y,\delta}(y, \delta) \ dyd\delta \]

\[ = \int_0^{\infty} \int_{-\infty}^{\infty} \gamma(y, \delta) \sum_{\delta_i \in \{\delta_i, \delta\}} \left\{ \left( \frac{\partial \ln f_{y,\delta}(y, \delta)}{dp} - \frac{\partial \ln f_{y,\delta}(y, \delta)}{dy} \frac{\partial \ln f_{y,\delta}(y, \delta)}{dy} f_{y,\delta}(y, \delta) \right) f_{y,\delta}(y, \delta) \ dyd\delta \]

\[ + \int_0^{\infty} \int_{-\infty}^{\infty} \gamma(y, \delta) \left\{ \frac{\partial \ln f_{y,\delta}(y, \delta)}{dp} - \frac{\varepsilon}{1 + p\varepsilon} \ln y \frac{\partial \ln f_{y,\delta}(y, \delta)}{dy} f_{y,\delta}(y, \delta) \right\} f_{y,\delta}(y, \delta) \ dyd\delta, \]

where the first equality uses the fact that \( f_{\delta|y}(\delta | y) = f_{\delta|y}(\delta | y) = 0 \), and the second equality is obtained similarly as the extensive elasticity term on government revenue above. \( \square \)