Labour Economics 71 (2021) 102004

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

Predicting the path of labor supply responses when state dependence matters☆

Zhiyang Jia, Trine Engh Vattø

Research department, Statistics Norway, Postboks 2633 St. Hanshaugen, 0131 Oslo Norway

ARTICLE INFO

JEL classification:
C35
C51
H24
H31
J22

Keywords:
Labor supply
Tax and transfer reforms
Gradual responses
State dependence
Microsimulation

ABSTRACT

The standard labor supply model ignores possible inertia originating from individuals' previous labor supply decisions and assumes immediate adjustments to policy reforms. In this study we develop a model where past labor market status have effects on present decisions: first, there is habit dependence in the taste for leisure; second, labor market opportunities reflect experiences of the previous period; and third, there is a disutility of deviating from the choice of last period (status quo). All these three components induce state dependence in labor supply behavior and gradual rather than immediate responses to tax and benefit reforms. The model is estimated with data of Norwegian females over the period 2003 - 2009. Simulation results from a tax rate change suggest that state dependence bring down the short-term (first-year) responses to one-third of the full effect, and the full effect is reached after about five years. Our results also suggest that the disutility of deviating from status quo, modeled as a fixed cost of switching, is the dominant driving force of sluggishness in labor supply responses.

1. Introduction

A number of studies suggest that long-term labor supply responses to policy changes are significantly larger than short-term responses (Kleven and Schultz, 2014; Gelber et al., 2020). Here, we would like to explore the underlying mechanisms behind such results. An improved understanding of what causes sluggish responses is crucial, as it will lead to better predictions of the path of labor supply responses to prospective policy reforms, and may also help policy-makers to design policies to reduce sluggishness in responses.

State dependence in labor supply has been discussed before, see Haan (2010); Hyslop (1999); Prowse (2012). Positive state dependence leads to sluggish response paths to policy changes and can be reflected by the differences between short- and long-term elasticities. As pointed out by Heckman (1981), past labor market experience has a causal effect on current labor supply behavior. The contribution of the present paper is to provide explanations to such patterns.

We develop a model that allows for different mechanisms that leads to sluggish labor supply responses to policy changes, extending the static one-period labor supply model of Dagsvik and Jia (2016). The components of sluggish responses are econometrically identified and used to simulate effects on the path of adjustment following from a tax policy change.

In our model, we let past experiences work on current choices via three different channels: first, we let past labor market experience directly affect current period preferences for leisure (Woittiez and Kapteyn, 1998; Kubin and Prinz, 2002); second, we allow past experience to influence the job opportunities faced by the individual, through the signaling and scaring effects (Spence, 1973; Arulampalam et al., 2000); and third, we allow for status quo bias in choice (Samuelson and Zeckhauser, 1988), reflecting people's tendency to avoid change. This third component can also be related to optimization frictions or adjustment costs, see Chetty (2012), Chetty et al. (2011), Kleven and Waseem (2013), and Gelber et al. (2020). Under suitable parametric assumptions, individuals' observed working hours over time can be used to separate the contributions of these channels, while taking account of unobserved individual heterogeneity.

Our model adds to a previous study by Haan (2010), who first introduced the idea of incorporating state dependence in labor supply de-

☆ We have received valuable comments on earlier versions of the paper from Thor O. Thoresen, John Dagsvik, Terje Skjerpen, Viktor Steiner, Frank Fossen, Peter Haan, Spencer Bastani and Regina Riphahn. We also thank participants at Wirtschaftspolitisches Seminar in Berlin, the IIPF conference in Taormina/Sicily, EALE in Turin, the Norwegian-German seminar in Munich, the IMA conference in Canberra, the IAAE conference in Milan, and seminar participants at Umeå School of Business and Economics for helpful discussions and comments.

E-mail addresses: Zhiyang.Jia@ssb.no (Z. Jia), Trine.Vatto@ssb.no (T.E. Vattø).

https://doi.org/10.1016/j.labeco.2021.102004
Received 18 April 2020; Received in revised form 15 February 2021; Accepted 14 May 2021
Available online 23 May 2021

cisions in a tax simulation setting.[1] We argue that disentangling the mechanisms of state dependence is important for improving the predictive power in a practical tax policy setting. More importantly, the different mechanisms of our model have different implications for individuals welfare, and call for different labor market policies to reduce potential welfare losses caused by sluggishness. Whereas the first component of state dependence refers to sluggishness in preferences and therefore is attributed to individual preferences, the two other components of state dependence may result from inefficiencies in the labor market.

We estimate the model on panel data of married or cohabiting females obtained from Norwegian administrative registers for the period 2003 - 2009. The model provides good predictions of observed labor market transitions over time. To further validate, we apply the estimated model to a holdout sample from the original data and find again that the model fits the data well. In addition, we show that the out-of-sample prediction performance of our model is on par with a much more flexible model where past labor market status enters "semi-parametrically".

We use the estimated model to simulate the effect of a tax cut and map out the time frame of labor supply adjustments. The simulation results suggest that there is an adjustment period of about five years before the full effect of the reform is approximately realized.[2] The first-year effect amounts to only about one third of the full effect. The predicted long-term effect is close to the predictions of a standard one-period model (without state dependence), as would be expected from the conceptual framework. We further conclude that status quo bias, which is modeled as a simple fixed switching cost in our setting, is the dominant source of sluggishness in responses to policy changes. To explore how strong the results depend on our parametric assumptions, we do a series of robustness checks where we re-estimate the model under alternative assumptions.

The paper is organized as follows: The model and its empirical specification are described in Section 2, whereas the data is described in Section 3. In Section 4 we discuss the estimated model and the simulation results. In particular, we describe how preference dependence, labor market constraint dependence and switching costs contribute to persistence, and present a simulated time frame of adjustment. Section 5 provides a conclusion.

2. The model

2.1. The standard one-period model

In the category of structural labor supply modeling approaches, the discrete choice model of labor supply based on the random utility modeling approach (Van Soest, 1995) stands out, as it has gained widespread popularity among public finance practitioners (Creedy and Kalb, 2005). This type of models can easily handle non-linear and possibly non-convex budget sets caused by taxation and are thus more practical than the traditional approaches based on marginal calculus. They are particularly useful to predict counterfactual labor supply effects of potential policy proposals to assist decision-makers. For example, the Norwegian government regularly uses a labor supply model to study the potential revenue effects of tax changes. The labor supply model is based on a particular type of discrete choice model denoted as the job choice model; see for example, Dagsvik and Jia (2016). According to this framework, labor supply decisions are viewed as the outcomes of individuals choosing among jobs, with additional constraints on the set of available jobs.

The job choice model is specified as follows: Each individual is assumed to have preferences within a set of 'jobs', where each market job (indexed by $z = 1, 2, \dots$) is characterized by disposable income $C(z)$, hours of work $h(z)$, and other non-pecuniary job attributes such as job-specific tasks to be performed, workplace locations and working environment quality. Disposable income for a given job is defined as $C(z) = f(h(z)w(z), I)$, where $w(z)$ is the offered wage rate for the given job z, I is non-labor income and $f(.)$ is the net-of-tax function. The offered wage rate $w(z)$ is assumed to be constant across jobs for a given individual.[3] The individual's utility of choosing job z is represented as $U(C, h, z)$, where the utility function is assumed to be additively separable, i.e. $U(C, h, z) = v(C, h) + \varepsilon(z)$.

The sets of available jobs from which the individuals choose, are individual-specific. Dagsvik and Jia (2016) show that it is sufficient to identify the model by introducing a measure of job opportunities representing the number of available jobs for a given working time option h, $m(h)$, where the number of non-working opportunities is normalized to one, i.e., $m(0) = 1$.

Hours of work for each job take a value within a given set H. It can be shown that applying the assumption of i.i.d. extreme value distributed error terms, $\varepsilon(z)$, the probability of a worker choosing one of the jobs with working time $h \in H$, can be written as,

$$\varphi(h) = \frac{m(h)exp(v(C(h), h))}{\sum_{x \in H} m(x)exp(v(C(x), x))} = \frac{exp(V(C(h), h))}{\sum_{x \in H} exp(V(C(x), x))}. \tag{1}$$

This expression is analogous to a multinomial logit model with payoff $V(C(h), h)$ in which the systematic part is the sum of the representative utility, $v(C, h)$, and the log of normalized number of available jobs for hours of work h, $m(h)$. Since $m(h)$ is not observable, it is estimated simultaneously with $v(C, h)$.

The microsimulation model based on the above framework provides guidance to Norwegian policy makers about the labor supply effects of prospective tax and transfer reforms. The results of the model are used to predict so-called "the day after" effects, which means that the time path of actual adjustment is neglected.

2.2. Extending the one-period model to allow for state dependence

The aim of the present study is to establish a model framework which accounts for the timing of labor supply responses. This requires us to extend the one-period model framework presented above to model labor supply decisions over time ($t = 1, \dots, T$). Following Haan (2010) and Prowse (2012), we assume individuals make labor supply decisions based on period-to-period optimizations.[4] We also assume, for the sake of simplicity, that state dependence forms a first order Markov chain over time, meaning that only the last period, and not the whole history of labor market outcomes, has a bearing on current decisions. In this setting, observed persistence in labor supply over time (on the individual level) is caused by both observed- and unobserved heterogeneity, as well as true state dependence, where past experience influences future decisions.

At first sight, introducing true state dependence seems to be straightforward in such a setting. One only needs to allow the periodic payoff V_t to depend on previous labor market experience in addition to current period's consumption and leisure. While there are many studies proposing potential theories of the underlying mechanisms, empirical models that explicitly specify how past labor experiences should be modeled are rare. The typical practice is to apply a "semi-parametric" setup, i.e., by including dummy variables for all possible pairs of transitions of labor market status in period $t-1$ and t. This setup is flexible, fits the data well and is capable of providing estimates of the causal effect of previous labor market status on current behavior. However as noted by

[1] See also Haan et al. (2015) and Haan and Uhlendorff (2013).

[2] In our context, the full effect is reached when the labor market reaches its equilibrium, i.e., changes in the distribution of working hours are negligible. In the following, we refer to this as the 'long-term' effect.

[3] Ideally, one would prefer to allow for unobserved heterogeneity across both jobs and workers. Unfortunately, identification of such a model is not guaranteed. See Dagsvik and Jia (2016) for a more detailed discussion of the specification of offered wage rates.

[4] We refrain from intertemporal substitution which can only be analyzed in a dynamic life-cycle model that allows for time inseparability (see e.g. Low and Meghir, 2017). This is beyond the scope of the present paper.

Haan (2010), not much can be learned about the underlying mechanisms. Moreover, since the number of unknown parameters increases quickly with the number of possible labor market states, the model can be quite complex.

There are two challenges which might explain why there are few empirical models of labor market dynamics that specify channels of the effects of past experience. First, the concept of state dependence is relatively broad, capturing several effects not precisely defined. Second, available data contain only limited information, which makes it challenging to disentangle the underlying mechanisms nonparametrically (Haan, 2010). While acknowledging these concerns, we think it can still be useful to make an attempt to develop a model that explicitly specify how past labor experience impact current decisions. This will inevitably require us to impose stronger assumptions. In turn, we will get a more parsimonious model which can shed more insights into the underlying mechanisms.

In the following, we assume that state dependence works through three different channels: preferences, labor market opportunities and disutility of deviating from last period's choice (status quo).[5] First, the value of leisure (time spent not working), or more precisely, the marginal rate of substitution between consumption and leisure, can be altered by past behavior. A number of studies suggest that individuals' preferences change over time by experience (Neuman et al., 2010; San Miguel et al., 2002). For example, individuals starting to work reduced hours, may experience that the additional time off is more valuable than previously thought, and place a higher value on the job options with shorter working time in the future. This is also consistent with habit formation in labor supply behavior (Woittiez and Kapteyn, 1998; Kubin and Prinz, 2002). We call this preference dependence and rewrite the systematic part of the payoff for working h hours in period t as $v(C(h), \kappa(h_{t-1})h)$, where $\kappa(h_{t-1})$ measures the effect of last period's labor supply decision on the taste of leisure.

Second, job opportunities may be altered by past labor market experience (see e.g. Heckman and Borjas, 1980). For instance, if employers consider previous non-participation as a result of depreciation of human capital (Pissarides, 1992) or as a signal of low productivity (Vishwanath, 1989; Lockwood, 1991), an individual that is inactive in the previous period may have less job opportunities in the current period than an otherwise identical worker. Using data from a field experiment, Eriksson and Rooth (2014) find evidence that employers attach a strong negative value on individuals who have contemporary unemployment spells last over 9 months. This supports the theory that employers may use information of past labor market behavior to sort job applicants. They also find that work experience can be seen as an important signal of productivity, which is consistent with the notion that work experience can be seen as a positive characteristic. We label this as labor market constraint dependence. A novel aspect of our framework (building on Dagsvik and Jia, 2016) is that latent labor market opportunities are explicitly modeled. This enables us to incorporate labor market constraint dependence by allowing the individual job choice set to depend on previous work experience. Within the framework of job choice, this is equivalent to allowing the number of jobs $m(.)$ to depend both on the hours of work alternative h and the hours of work choice in the previous period h_{t-1}. Thus, we have, $m = m(h, h_{t-1})$.

Third, individuals have a tendency to stick to their previous choices, which is labeled as "status quo bias" by Samuelson and Zeckhauser (1988). While many consider this as an indication of irrational decision-making, it can be explained by material or mental costs in connection to changes in behavior, resulting from the presence of actual adjustment costs such as informational costs (Matejka and McKay, 2014; Steiner et al., 2017) or job search cost (Hyslop, 1999), cognitive misperceptions such as loss aversion (Dunn, 1996), and anchoring (Furnham and Boo, 2011). In this paper, we model the status quo bias as a disutility of deviating from the previous choice. In other words, individuals will need to pay a "switching cost" if they choose a labor supply alternative other than last period's choice. The concept of switching cost is similar to the adjustment costs in the optimization friction models (Chetty et al., 2011; Gelber et al., 2020). The term "switching cost" is borrowed from marketing science, where consumers have costs associated with switching between the products of competing firms (see e.g., Dubé et al., 2010). In our context this cost will depend on both the previous job and the current job and we denote it as $S(h, h_{t-1})$. In addition, we assume that it enters the utility additively, i.e. the actual utility obtained by choosing hours of work, after all three channels are considered, can be written as $\tilde{U}(C, h \mid h_{t-1}) = v(C(h), \kappa(h_{t-1})h) - S(h, h_{t-1}) + \varepsilon(h)$. Under the assumption that $\varepsilon(h)$ is i.i.d. extreme value distributed, we can show that the labor supply probability in period t given hours in previous period h_{t-1}, can be written as,

$$P(h \mid h_{t-1}) = \frac{m(h, h_{t-1})exp(v(C(h), \kappa(h_{t-1})h) - S(h, h_{t-1}))}{\sum_{x \in H} m(x, h_{t-1})exp(v(C(x), \kappa(h_{t-1})x) - S(h, h_{t-1}))}. = \frac{exp(\widetilde{V}(C(h), h, h_{t-1}))}{\sum_{x \in H} \widetilde{V}(C(x), x, h_{t-1})}. \quad (2)$$

where $\widetilde{V}(C(h), h, h_{t-1}) = v(C(h), \kappa(h_{t-1})h) - S(h, h_{t-1}) + ln(m(h, h_{t-1}))$.

Note that the normalized number of jobs $m(\cdot)$ and switching cost $S(\cdot)$ are not observed so they need to be estimated together with the periodic utility $v(\cdot)$. Unfortunately, similar to the case for the static job choice model, the extended model is not non-parametrically identified. To be precise, one is able to non-parametrically identify $\widetilde{V}(C(h), h, h_{t-1})$ up to an additive constant, but not able to separately identify $v(\cdot)$, $m(\cdot)$ and $S(\cdot)$. This implies that while we can obtain the total effect of state dependence $\frac{\partial \widetilde{V}(C,h,h_{t-1})}{\partial h_{t-1}}$[6], we cannot non-parametrically disentangle effects via the three different channels mentioned above.

To obtain full identification, we need either obtain more information or impose stronger assumptions. In principle, stated preference data might be helpful in this situation. For example, if we can obtain information on "desired" hours of work, namely, hours of work when there is no restriction on labor demand and switching cost, we would be able to identify $v(\cdot)$. See for example Euwals and Van Soest (1999) and Bloemen (2008) for studies using desired hours of work to help with identification of preferences. Once individual utility $v(\cdot)$ is identified, the remaining job is just to distinguish $m(\cdot)$ and $S(\cdot)$. This can be done if there is a shock that impact only one of these two factors, for example, a demand shock which reduces the demand for labor but not the switching cost. However, going down this path is not straightforward. The reason is twofold: firstly it is questionable that the "desired" hours of work obtained in the survey actually represents a choice purely generated by preferences, and secondly it is often difficult to find a shock that impact only one but not the other factor, and moreover, shocks typically impact different individuals differently, which could practically make identification impossible.

Some exogenous variations can be quite helpful in terms of identifying the partial effects. For example, changes in the tax system can generate variations in disposable income $C(h)$. When wage rate and hours of work are kept constant, this will help identifying the partial effect of consumption on utility. The exclusion restriction that non-labor income does not affect $m(\cdot)$ and $S(\cdot)$ and impacts utility via consumption only, has a similar effect as a change in the tax system. On the other hand, $m(\cdot)$ can be seen as a proxy for labor demand. We expect that regional and time variation in the labor market tightness would have effect on individual's behavior only through $m(\cdot)$, which will help us identifying the partial effect of $m(\cdot)$ from $v(.)$ and $S(.)$.

[5] Note that state dependence can also work through wages. In a robustness check we allow for state dependence in wages, but we do not find that this affect our results, so we abstract from this channel in our baseline model.

[6] Strictly speaking, we should not use the partial derivative notation here since the hours of work is discrete.

However, none of these exclusion restrictions ensure identification as we still need additional assumptions to pin down the levels, see Theorem 2 in Dagsvik and Jia (2016) for similar arguments. In our empirical analysis presented below, identification is achieved by imposing parametric assumptions. This raises the concern that the results we obtain are products of the assumptions we make, and do not reflect any "deep parameters" or the underlying mechanisms. To deal with this concern, we do extensive robustness tests to check how sensitive our main results are to sensible changes in the empirical specifications. Although robustness of results is never a guarantee, it certainly makes our analysis more persuasive and trustworthy.

2.3. Unobserved heterogeneity and the likelihood function

It is well-known that not only observed heterogeneity (wages, age, children, education), but also persistence in unobserved individual heterogeneity leads to spurious state dependence (Heckman, 1981). Following Haan (2010), we model persistent unobserved heterogeneity in a latent class framework. We assume that individuals can be classified into K different (unobserved) types, for which some key parameters differ. An individual's type is assumed to be the same over time. The fraction of individuals of type $k = (1, ., K)$ is estimated within the model by $\pi_k \in (0, 1)$. This leads essentially to a finite mixture model and has the advantage that unobserved heterogeneity can be handled flexibly, without imposing a parametric structure.

Persistent unobserved individual heterogeneity creates a pattern of serial correlation in the error terms of the utility function. We abstract from a more general structure of serial correlation or the related concept of habit persistence (as defined by Heckman, 1981). A number of empirical studies seem to find that serial correlation and habit persistence are of less importance than state dependence (Hyslop, 1999; Prowse, 2012; Seetharaman, 2004). We therefore do not expect that excluding these remaining aspects of persistence is crucial to our interpretation of the results.

The likelihood of an individual choosing the sequence $(h_1, h_2 \cdots h_T)$ can be written

$$\varphi(h_1, h_2 \cdots h_T) = \sum_k \pi_k \prod_t P_k(h_t|h_{t-1}) \tag{3}$$

where the conditional probabilities $P_k(h_t|h_{t-1})$ for each unobserved type are given by Eq. (2).

Note that as individuals' decision at period t is conditional on decision at period $t-1$, the choice at the initial period ($t = 1$) depends on unobserved behavior at $t = 0$ and cannot be treated as random. In order to solve this problem of initial conditions, we apply the method suggested by Wooldridge (2005). This method has been applied in several studies of labor market dynamics, including Haan (2010), Prowse (2012) and Haan et al. (2015).

Keep in mind that the model assumes that individuals only consider the present utility when deciding whether to adjust working hours or not, similarly to in the main analysis of Gelber et al. (2020). In reality individuals may also pay attention to the future discounted utility. Thus, when switching costs and state dependence are fitted to the observed pattern of persistence, the estimates may be affected by this modeling choice. However, for a relative comparison of the different components of state dependence and for practical simulations, this should not be crucial.

2.4. Empirical specifications

In the following, we present the specification of the main model and discuss which alternative specifications that can be relevant and should be addressed in the robustness checks.

Hours of work and wage regression

For all individuals, we discretize the information on working time by dividing it into five categories based on weekly hours of work: $h \in \{0,\ 1-19,\ 20-34,\ 35-40,\ 41+\}$. Thus, we assume that each individual chooses a job characterized by one of these five working hour options in each time period. Individual wages are obtained from a Heckman wage regression; see Table A.1 in the Appendix. A tax simulator is used to compute taxes for each option.

The periodic utility function

In accordance with Dagsvik and Jia (2016), we assume that the deterministic part of the utility function can be represented by a Box-Cox function,

$$v\left(C_t, h_t \,|\, h_{t-1}\right) = \alpha_0 \frac{\left(C_t - C_0\right)^{\alpha_1} - 1}{\alpha_1} + \beta_0 \frac{\left(\bar{h} - h_t\right)^{\beta_1} - 1}{\beta_1} \tag{4}$$

where C_t is disposable income expressed as $C_t = f_t(h_t w_t, I_t)$. C_0 represents the minimum or subsistence household-adjusted consumption level, set here to NOK 50,000 (about USD 8,500 or EUR 6,200). $\bar{h}$ is defined as 80 hours per week, such that $(\bar{h} - h_t)$ measures leisure time. As in the standard one-period model, leisure and consumption preferences are revealed by observed choices. In the extended model framework presented here, we model state dependence in preference by allowing for the taste of leisure, β_0, to depend on the previous working time decision, h_{t-1}.[7] We let β_0 be specified as follows,

$$\beta_0 = b_0 + \mathbf{b}_1^{'}\mathbf{x_t} + \mathbf{b}_2^{'}\mathbf{I}\left(h_{t-1}\right) + \mathbf{b}_3^{'}\mathbf{I}\left(h_0\right) \tag{5}$$

where $\mathbf{x_t}$ is a vector of observed individual characteristics (age and number of children), $\mathbf{I}(h_{t-1})$, is a unit vector of the individual's previous working time, and $\mathbf{I}(h_0)$, is a unit vector of the initial working time.[8] It follows that if $\mathbf{b}_2' \neq \mathbf{0}$, then the previous working time decision affects current leisure preferences, and thus represents what we refer to as preference dependence.

Regarding the choice of utility function another alternative is the polynomial (Van Soest et al., 2002). It has the advantage of being flexible and easy to estimate because it is linear in parameters. However, in contrast to the Box-Cox setup, this specialization is not guaranteed to be quasi-concave and monotone in consumption. To study whether our results are robust to the choice of different functional forms of the utility, we report results from a version in which the deterministic part is a quadratic utility function.

The opportunity measure

The job opportunity measure, $m(h_t)$, is considered to be a sufficient statistics for the choice sets of available jobs and represent labor market restrictions (Dagsvik and Jia, 2016). It can be seen as belonging to the demand side of the labor market in combination with labor market regulations.[9] Without loss of generality, let

$$\theta = \sum_{h>0} m(h) \quad \text{and} \quad g(h) = m(h)/\theta,$$

where one can interpret θ as the normalized total number of jobs (relative to non-participation) and $g(h)$ as the fraction of jobs available to the agent with offered hours of work equal to h. In empirical studies, θ is often assumed to be a function of individual characteristics, such as education, while $g(h)$ is independent of individual characteristics, since hours restrictions are considered to be determined to a large extent by

[7] We have tested an alternative specification where we allow that taste for consumption, α_0, also depend on previous labor market behavior. Unfortunately, we were not able to achieve convergence due to numerical problems.

[8] Wooldridge (2005) suggests estimating unobserved heterogeneity conditionally on initial period observations, as described in Section 2.3. A similar strategy for the structure of state dependence and initial conditions can be found in Haan et al. (2015).

[9] There is a fairly strong degree of labor market regulations in Norway concerning working time, described in the law on labor relations ("Arbeidsmiljøloven").

labor market institutional regulations and negotiations between unions of employers and workers (Dagsvik et al., 2014; Dagsvik and Jia, 2016). Following similar arguments, we assume that past labor market experience will impact θ, but not $g(h)$:

$$m(h, h_{t-1}) = \theta(h_{t-1})g(h). \tag{6}$$

Thus, state dependence and individual heterogeneity in job opportunities are accounted for by the total number of jobs available (relative to non-participation). To be precise, we specify $\theta_t = \theta_t(h_{t-1})$ as,

$$ln(\frac{1}{\theta_t(h_{t-1})}) = \gamma_0 + \gamma_{11}(V_r/U_r)_t + \gamma_{12} edu_t + \boldsymbol{\gamma_2'}\mathbf{I}(h_{t-1}) + \boldsymbol{\gamma_3'}\mathbf{I}(h_0) \tag{7}$$

where V_r/U_r refers to the vacancy to unemployed ratio which is a measure of regional labor market tightness, and edu_t symbolizes the individual's education level. We allow for that current labor market opportunities depend on both the extensive and the intensive margin decisions of the previous period, $\boldsymbol{\gamma_2'}\mathbf{I}(h_{t-1})$. We let $g(h)$ be uniformly distributed among working time options, except for a possible peak (estimated within the model) for full-time jobs. This essentially means that we assume there are equal numbers of short and long part-time jobs to choose from, and a larger number of full-time jobs available. Note that this specification of the opportunity measure is equivalent to introducing suitable dummy variables at the full time peak in the utility specification of the conventional discrete choice specification, see for example Van Soest (1995) and Creedy and Kalb (2005). In the robustness checks below, we allow for an alternative specification of $g()$. Note that the above assumption implies that the specified state dependence via opportunities will only impact the extensive margin but not the intensive margin, since past labor market status only impact the total number of available market jobs but not the number of options of different working hours.

The switching cost

Lastly, we specify switching costs, $S(h_t, h_{t-1})$, which represents a disutility caused by a change in working time from period $t-1$ to t. The most flexible specification would be to allow the cost to differ across all possible working time transitions. This will be equivalent to including one dummy variable for each possible pair of transition in hours of work. Such a setting essentially leads to a "semi-parametric" model, as mentioned above, which would make it impossible to distinguish switching costs from the other channels of state dependence. Instead, we impose a minimalist assumption and define switching costs as a constant term for the disutility associated with deviating from the previous labor supply decision, namely.

$$S(h_t, h_{t-1}) = \begin{cases} s & \text{if } h_t \neq h_{t-1} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

where s is a constant. Although switching cost may be related to preferences, it does not affect the individual's marginal rate of substitution between consumption and leisure in our framework, as is the case for preference dependence. This simple structure of switching costs allows us to identify it separately from preference dependence. Preference dependence is defined as an upward or downward shift in the smooth function of leisure preferences (Eqs. 4 and 5), whereas switching costs are characterized as a fixed cost of altering working time from last period's decision (Eq. 8). This assumption is obviously rather restrictive. However, as we will show below, the model is capable to replicate the observed labor market transitions both in- and out-of-sample. In particular, the out-of-sample prediction performance of the model is on par with a much more flexible model where past labor market status enters "semi-parametrically". This may suggest that our specification is a reasonable approximation to the "true" underlying mechanism.

3. Data

3.1. Data and summary statistics

The model is applied to data from merged administrative registers of Statistics Norway, which contain detailed information on household composition, reported income and socioeconomic characteristics, as well as monthly wages and working time for the majority of Norwegian wage earners in the period 2003 - 2009. We limit the data set to married and cohabiting women aged 25 - 62 years, and exclude the self-employed, disabled persons and students. To abstract from the effects of various welfare transfers, we limit the sample to women whose partners' total pre-tax income level exceeded NOK 150,000 (about USD 25,500 or EUR 19,000) per year. As non-participation is a result of choice in our model framework, we further exclude individuals who are recipients of unemployment benefits.[10] The remaining balanced data set consists of about 240,000 observations each year.[11]

Information on working time is available for about 70 percent of the observations each year, and is based on employer's reports and administrative registers. To avoid attrition and selection effects, we impute the missing working time information (see Appendix).

For practical reasons we select a random ten percent sample (the estimation sample) for use in the estimations and we use another three percent sample (the validation sample) to check how the estimated model perform out of sample. Table 1 provides an overview of the main characteristics of the estimation sample. Over the period 2003 to 2009, on average about 6 percent are not labor market participants, 19 percent work short part-time, 31 percent work long part-time, 37 percent work full-time and 8 percent work overtime.

3.2. Observed persistence in specific working times

Transition of hours of work over time can be presented in a Markov transition matrix. It provides information about the fraction of individuals starting out in a certain working time category who remain there for the next period, and how many switch to other working time categories. The diagonal elements of the transition matrix can be interpreted as measures of persistence. The higher the magnitude, the larger the fraction of individuals who have the same labor status as last period, and thus the stronger the persistence.

In order to also take account of the inflow into a specific labor market choice,[12] persistence can be measured in the form of stability, as the number of individuals in a specific working time category in both periods divided by the number of individuals present in the working time category in at least one of the two periods.[13] This persistence rate takes a value between 0 and 1, where a higher value implies higher persistence

[10] Recipients of unemployment benefits can be regarded as involuntary non-participants in the labor market as they must register as active job seekers and be willing to take any job anywhere in Norway at short notice. As a robustness check, we include individuals with unemployment benefits in the sample, and use their wage income (excluding unemployment benefits) to assign working time. This does not affect our results significantly. The fraction of non-participants is actually slightly reduced, as individuals typically receive unemployment benefits over only a short period of time.

[11] When restricting the panel to be balanced we lose a number of individuals who only fulfill the requirements in some but not all years. Only about 30 percent of individuals observed for at least one year are observed in all periods (2003 - 2009). However, this does not seem to significantly affect either the observed characteristics of the sample or the observed choice probabilities.

[12] For instance, labor market constraint dependence may reduce the outflow from non-participation to participation, but may also reduce the inflow from participation to non-participation, because individuals participating the previous period obtain more job opportunities and are therefore more likely to participate next period as well.

[13] This type of measure has, for example, been frequently used in the literature on school mobility, see e.g., Dobson et al. (2000).

Table 1
Characteristics of the pooled sample 2004 - 2009.

Variable	Mean	Std. Dev.	P25	Median	P75
Imputed wage rate (NOK)	167	35.9	145	160	196
Non-labor income (NOK 1,000)	28.4	228	2.8	18.2	30.7
Partner's gross income (NOK 1,000)	584	1425	367	459	618
Age	45.3	8.4	39	45	52
Child(ren) under age 6	0.26	0.44	0	0	1
Child(ren) under age 12	0.45	0.50	0	0	1
Low education (≤10 years)	0.38	0.49	0	0	1
Regional vacancy-to-unemployed ratio	0.36	0.24	0.16	0.32	0.46
Number of individuals			23,679		

Notes: The sample consists of married women. Income and wage rates are adjusted to 2007-NOK (NOK 1≈USD 0.17≈EUR 0.13). P25 and P75 refer to percentile 25 and percentile 75, respectively.

Table 2
Observed annual transition probabilities and persistence rates in the period 2003 - 2009.

	Year t+1				
Year t	None	Short p-t	Long p-t	Full-time	Overtime
None	**83.9%**	13.8%	1.3%	0.9%	0.2%
Short p-t	2.8%	**76.2%**	16.3%	4.0%	0.8%
Long p-t	0.1%	7.5%	**77.2%**	13.1%	2.2%
Full-time	0.0%	1.7%	9.5%	**80.8%**	8.0%
Overtime	0.0%	1.5%	6.4%	34.4%	**57.6%**
Persistence rate	**0.77**	**0.63**	**0.63**	**0.67**	**0.40**

The observed annual transitions of the sample of married women by working time categories. The allocation into working time categories is based on reported working hours per week: 0 ("None"), 1–19 ("Short part-time"), 20–34 ("Long part-time"), 35–40 ("Full-time") and 41+ ("Overtime"). The persistence rate is defined as the number of individuals in the specific working time arrangement in both periods divided by the number of individuals present in the working time arrangement in at least one of the two periods.

and less mobility across categories. If no one chooses the same category two periods in a row, the persistence rate equals 0, whereas if everyone chooses a given category repeatedly, the persistence rate equals 1.

Table 2 presents the observed annual transition probabilities in the period 2003 - 2009 and the average persistence rates for each working time category, as defined above. We see strong persistence in observed labor supply decisions. The persistence rate is highest for non-participation (0.77) and lowest for overtime work (0.40). The aggregate intensive margin persistence rate, is a weighted average of the persistence rate for each state and is slightly lower than that of the extensive margin, with a value of 0.62. In the next section, we apply these persistence measures to evaluate how the different components of state dependence affects persistence.

4. Results

In the following we describe the main estimation results, including a discussion of the estimated parameters and how well the model fits data, both using the estimation sample and a holdout validation sample. Next, we show how the estimated model is used to simulate the path of labor supply adjustment, and how each channel of state dependence contributes to the sluggish response path. Finally, we demonstrate that our main results are robust to alternative parametric assumptions and discuss the policy relevance of our results.

4.1. Estimated parameters

The model is estimated by the method of maximum likelihood, where the likelihood function is given by Eq. (3). $t = 0$ refers to the initial conditions in year 2003 and $t = T$ refers to year 2009. The parameters for the utility function, job opportunity measure, switching costs and probabilities of each unobserved type of individual are estimated simultaneously. For each individual, the observed path of labor market decisions is given by the sequence of working time (five categories) over the six-year period (2004 - 2009), such that $5^6 = 15,625$ different working time paths are possible.[14] The estimated parameters are reported in Table 3.

The utility function turns out to be concave and to increase with respect to consumption and leisure for all individuals, and the job opportunity measure has the expected sign. We find evidence of individual heterogeneity in preferences and job opportunities. In particular, we find observed heterogeneity in preferences related to age and the presence of children, and in terms of unobserved heterogeneity in the job opportunity measure.

Significant negative estimates of $\mathbf{b}_2' = (b_{22}, b_{23}, b_{24}, b_{25})$ imply positive preference dependence, as previous working time experience reduces the subjective disutility of working the present period (or equivalently reduces the value of leisure).

Negative estimates of $\boldsymbol{\gamma}_2' = (\gamma_{22}, \gamma_{23}, \gamma_{24}, \gamma_{25})$ imply positive labor market constraint dependence, as individuals working in the previous period obtain a larger job opportunity measure, and are thereby more likely to participate in the labor market in the current period. Conversely, non-participation in the previous period implies reduced labor market opportunities in the current period, leading to lags in the transition from non-participation to participation. The size of the effects depends on the intensive margin (hours of work) in the previous period; working short part time seems to have a limited effect on job opportunities (γ_{22}), while the other three working options have larger effects that are of a similar magnitude ($\gamma_{23}, \gamma_{24}, \gamma_{25}$). Thus, working short part-time does not help to improve job opportunities and relax labor market constraints in the future as effectively as the other three working time options.

The switching cost estimate, s, is positive and significant. This suggests that deviating from the previous period's labor supply decision induces a loss in utility. In order to obtain a rough idea of the magnitudes involved, we compare the estimated utility loss of switching costs to the estimated utility loss of a reduction in disposable income. Simple calculations for a representative agent show that the estimated switching costs equal a disposable income reduction of about NOK 40,000 (USD 6800 or EUR 5,000). For a full-time working female this amounts to 1–2 monthly salaries, which suggests that switching costs are quite substantial.

4.2. Goodness of fit

Given the random utility framework, our model does not predict choice directly, but rather choice probabilities. We follow a commonly used method to evaluate the model' s predictive ability by comparing

[14] Data for year 2003 are used for initial conditions only.

the aggregated predicted choice probabilities against the observed share of individuals.[15]

First, we check how our model predict the observed empirical patterns on the estimation data set (in-sample tests).[16] We find that the model performs well in terms of the marginal distribution of choices which is a standard check in the one-period model. More interesting for the focus of the present paper is to evaluate whether the model can reproduce the observed pattern of transitions in the data. We find that the model reproduces the observed one-year transitions of the estimation sample very well, see Table A.2 in the Appendix, which can be compared to its observed counterparts in Table 2. The predicted persistence rates are also similar to the observed persistence rates as defined in Section 3. We further find that the model fits to a large extent the observed heterogeneity in the year-to-year transitions matrix between low and high wage earners, presented in Table A.3 and A.4 in the Appendix.

Second, following the method of holdout validation, we pick an independent subsample of our original data source to perform out-of-sample goodness of fit tests, as it has been argued that independent data sets for estimation and validation will better allow for external validation of the underlying behavior model than in-sample predictions (Schorfheide and Wolpin, 2016; Parady et al., 2021).[17] Moreover, the dynamics within the model are Markovian, so testing the fit over a longer period than one-year is a meaningful exercise to check the performance of the model. This is done by conditioning on the first period's observed behavior and then letting the model predict subsequent labor supply behavior over five years. We find that the predicted five-year transition matrix captures in large parts the observed counterparts, see Table A.5 in the Appendix.[18]

For comparison, we also estimate a "semi-parametric" model similar to that of Haan (2010), where lagged dependent variables enter as alternative specific effects. Such "flexible" models have more parameters and less restrictions, and it can therefore be expected that they perform better in-sample goodness of fit measures. However, our model is more ambitious in the sense that we explicit model and distinguish between the different mechanisms of state dependence. We find it reassuring that our model performs similarly to the more flexible model in the out-of-sample predictions checks summarized in Table A.5–A.7 in the Appendix.

Table 3
Baseline model coefficients.

	Parameter	Coefficient	Std. error
Probability distribution $(\alpha_0, b_0, \gamma_0)$			
Probability, unobserved type 1	p_1	0.5569***	0.0219
Probability, unobserved type 2	p_2	0.3648***	0.0204
Probability, unobserved type 3	p_3	0.0783***	0.0036
Preferences, consumption			
Constant (scale 10^{-4})	α_0	0.9579***	0.0412
Exponent	α_1	0.7277***	0.0126
Preferences, leisure			
Constant (scale 1/80), unob. type 1	b_{01}	4.9096***	0.2299
Constant (scale 1/80), unob. type 2	b_{02}	4.3138***	0.2183
Constant (scale 1/80), unob. type 3	b_{03}	3.4534***	0.1993
Exponent	β_1	-3.5190***	0.0570
Taste modifiers			
Age (scale 10^{-1})	b_{11}	-0.6023***	0.0669
Age squared (scale 10^{-2})	b_{12}	0.0765***	0.0074
Child(ren) under age 6	b_{13}	0.1118***	0.0144
Child(ren) under age 12	b_{14}	-0.0627***	0.0135
Preference dependence			
Short part-time, period t-1	b_{22}	-1.5328***	0.0863
Long part-time, period t-1	b_{23}	-2.0553***	0.0933
Full-time, period t-1	b_{24}	-2.3131***	0.0989
Overtime, period t-1	b_{25}	-1.9576***	0.0936
Switching costs			
Constant	s	1.6055***	0.0094
Opportunity measure (inverse)			
Constant, unob. type 1	γ_{01}	0.1894*	0.1915
Constant, unob. type 2	γ_{02}	4.0794***	0.1816
Constant, unob. type 3	γ_{03}	-0.0101	0.1642
Labor market tightness	γ_{11}	-0.6892***	0.1121
Low education ($\leq$ 10 years)	γ_{12}	0.0807	0.0632
Lab. market constraint dependence			
Short part-time, period t-1	γ_{22}	-0.0666	0.0876
Long part-time, period t-1	γ_{23}	-2.7181***	0.1919
Full-time, period t-1	γ_{24}	-2.9191***	0.2736
Overtime, period t-1	γ_{25}	-3.4037***	0.3724
Opportunity density			
Full-time peak	$g(h_t)$	0.5534***	0.0111

Notes: The estimation sample of married women contains 23,679 individuals. Initial working times are included (estimates are suppressed) such that unobserved heterogeneity is estimated conditionally on the initial period observations. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

4.3. Simulating the adjustment path

We now use the estimated model to simulate the effect of a tax cut on the path of labor supply adjustment. We describe the adjustment path for a permanent cut in the general tax rate. We also compare the results to simulations from a static version of the model, without state dependence. The static model is estimated on the same data, pooled for the period 2004 - 2009.

We describe the adjustment path by means of developments in labor supply elasticities. The elasticities are obtained by simulating the average predicted working time across individuals before and after a general tax cut applies. They can be interpreted as the percentage change in mean working hours when the net wage is increased by 1 percent. The labor supply elasticity is decomposed into a participation elasticity and an elasticity conditional on participation, which measure the extensive and intensive margins, respectively.

In order to estimate labor supply elasticities over time, we start from the observed initial labor supply choices, and let the model simulate the labor supply decisions for subsequent years according to the formula $\sum_k w_k \prod_t \varphi_k(h_t|h_{t-1})$, where w_k is the individual weight of unobserved type k. We compare a reference path with no tax cuts to an alternative path reflecting a permanent tax cut. Individual characteristics, including non-labor income and gross wage rate, as well as the tax schedule (apart from the tax cut) are kept constant in the simulations. Thus, the labor supply elasticities reflect the effect of the tax cut only.

Panel A in Fig. 1 provides a graphical illustration of the results. The intensive margin responses are of a larger magnitude than the extensive margin responses, but the transitions over time follow relatively similar

[15] Although this method does not provide a clear quantitative evaluation of the model performance, it provides simple and intuitive outcome comparisons. Quantitative goodness of fit performance measures proposed for discrete choice models in the literature range from general measures (the sum of squared error (SSE) and alike), likelihood based indexes (MacFadden R^2) to more specific measures designed for discrete choice model (percentage of correct predictions), see Parady et al. (2021) for an extensive review.

[16] To facilitate model predictions, we use the so-called empirical Bayes-method (Skrondal and Rabe-Hesketh, 2004) to specify individual-specific probabilities (the posterior distributions) of belonging to each unobserved type, $k = (1, 2, 3)$. The individual weights for each type are defined as $w_k = p_k \prod_t \varphi_k(h_t|h_{t-1}) / \sum_{j\in\{1,2,3\}} p_j \prod_t \varphi_j(h_t|h_{t-1})$.

[17] We have also performed an alternative out-of-sample test in which we have re-estimated the model over a shorter period of time (2003 - 2006), and then used this estimated model to compare simulated and observed outcome in-sample (2004 - 2006) and out-of-sample (2007 - 2009). The results are similar to the main model, and it seems to perform well in terms of out-of-sample goodness of fit also in this case.

[18] The predicted five-year transition matrix is not as accurate match to its observed counterparts as the one-year predictions. This can be expected as uncertainties will accumulate over time and impact the precision of a multi-period transition matrix (Haan, 2010).

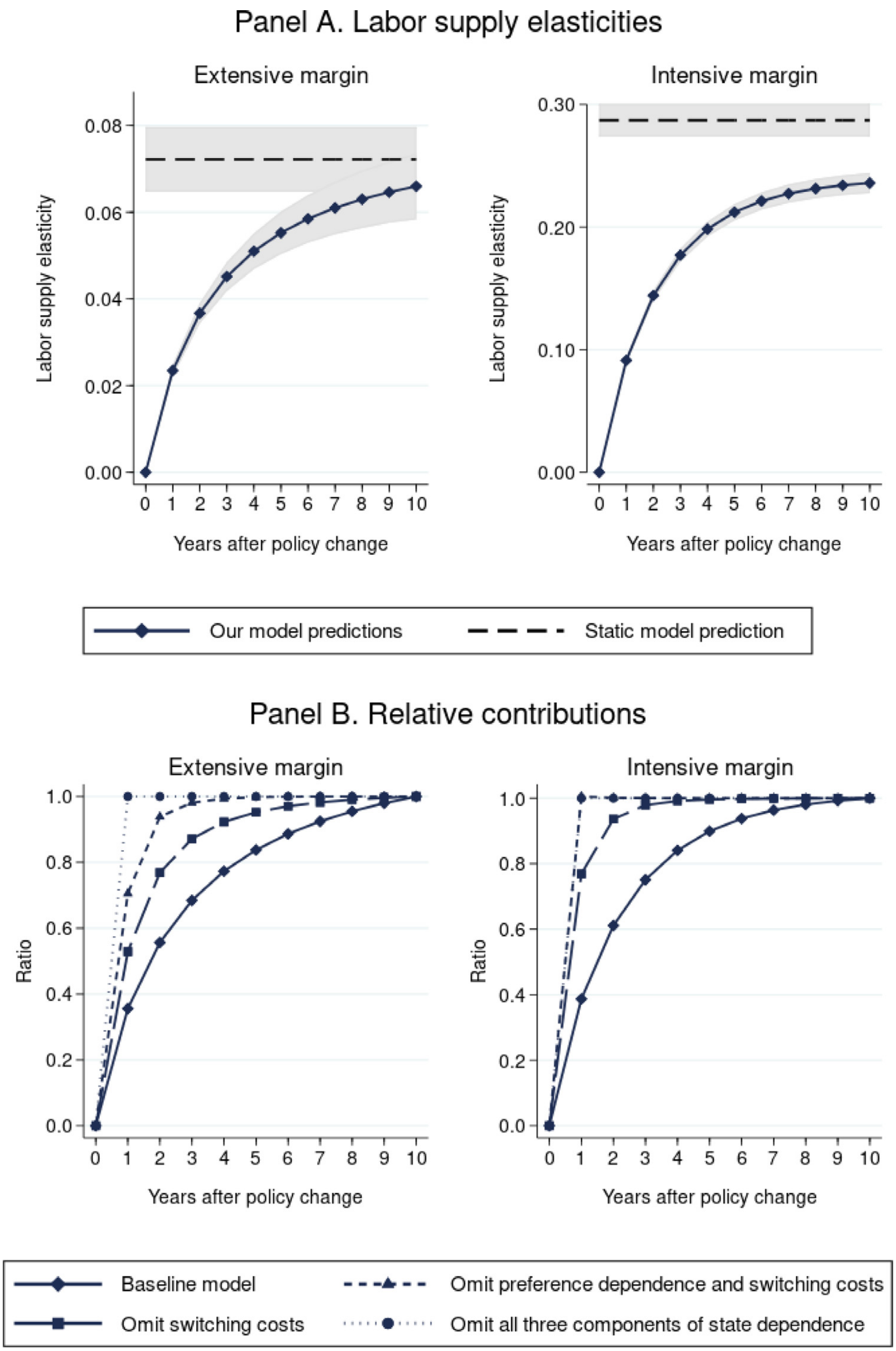

Fig. 1. Simulated labor supply adjustment and the relative contributions Notes: Panel A. The simulated labor supply responses to a general tax rate cut. The labor supply elasticities refer to the percentage change in participation with respect to the net wage rate ("extensive margin") and hours of work conditional on working with respect to the net wage rate (intensive margin), respectively. The shaded area represents the 95 percent confidence interval obtained by non-parametric bootstrapping, 50 repetitions. Panel B. Switching costs are omitted by replacing $\hat{s}$ (estimated parameters reported in Table 3) by 0. Preference dependence is omitted by replacing $\hat{b_{2j}}$ by a constant average, $\bar{b_2}$. Labor market constraint dependence is omitted by replacing $\hat{\gamma_{2j}}$ by a constant average, $\bar{\gamma}_2$. The 10th-year elasticity is normalized to 1.

paths.[19] At the extensive margins, the first-year responses of a permanent tax rate change account for one-third of the long-term responses, which are achieved after approximately five years.[20] At the intensive margin, the first-year responses account for slightly more than one-third of the long-term responses, which are achieved after about five years. As with the results of Haan (2010), this suggests a rather slow adjustment speed.[21]

When comparing the elasticities obtained from our model to the prediction of a one-period model without state dependence, we see that the difference is especially pronounced in the short-term. This suggests that

[19] In general, one often finds that extensive margin responses are larger than intensive margin responses (see e.g. Heckman, 1993). However, elasticities depend on characteristics of the data sample due to non-linearities. We expect that high participation rates in the Norwegian context contribute to the modest elasticities we find at the extensive margin.

[20] We define the 10th year response as the long-term elasticity, and report the number of years until at least 90 percent of the long-term effect is reached.

[21] Gelber et al. (2020), on the other hand, concludes from an analysis of income responses to changes in the social security earnings test in the United States that adjustment is completed after about three years.

state dependence attenuate responses in the short-term, but eventually reach predictions from the static model in the long-term (although there is a small gap at the intensive margin), which suggests that state dependence do not prevent individuals from adjusting in the long-term.[22] This allows us to establish another simple measure of the adjustment speed implied by the model, which is the short-vs-long term elasticity ratio. The smaller the ratio is, the slower the responses will be, and the longer the full effect takes to be realized. The ratio based on our model is 0.36 and 0.39 for the extensive and the intensive margin respectively.

According to Gelber et al. (2020) it has over time been postulated that long-term responses are significantly larger than the short-term responses, due to frictions that impede adjustment in the short term. This is supported by a number of reduced form studies finding a gradual adjustment of employment to a policy change, although the methodology does not give much guidance on how to interpret the time path of the treatment effect. An alternative to the methods described in Section 4.2 to test our model's performance is thus to compare model predictions by means of quasi-experimental findings on responses to tax reforms.[23] Although a complete approach along these lines is not within the scope of the present paper, it can be noted that Vattø (2020) uses quasi-experimental estimation over the same tax reform period to conclude that the long-term elasticity is reached after about five years, and accounts for almost twice the size of the short-term elasticity. A slow adjustment path is also supported by other studies of earnings responses evaluating tax reforms where long-term responses are found to be more pronounced than short-term responses, see e.g., Giertz (2010), Bækgaard (2014), Kleven and Schultz (2014), Neisser (2017), Jongen and Stoel (2019) and Gelber et al. (2020).[24]

4.4. How the different components of state dependence contributes

Model estimates suggest that all three components contribute to true state dependence. However, standard statistical tests on the parameter estimates cannot provide useful information on the relative importance of these three components. We therefore suggest an alternative approach where we try to attribute the degree of sluggishness to each component. To do this, we compare the simulated persistence rate (defined in Section 3.2) and the short-vs-long term elasticity ratio (defined in Section 4.3) obtained from three constrained models, where we leave out one component of state dependence or switching costs from the full baseline model.

To be precise, we use the estimated baseline model to perform three simulations in which we impose the following restrictions: (i) Past labor supply decisions have no impact on leisure preferences, i.e., b_{2j} is replaced by a constant average, $\bar{b_2}$; (ii) Past labor supply decisions have no impact on labor market constraints, i.e., γ_{2j} is replaced by a constant average, $\bar{\gamma_2}$; (iii) Past labor supply decisions have no impact on switching costs, i.e., s is replaced by 0.

In Table 4 we report the persistence rate and the short-vs-long term elasticity ratio at the extensive and intensive margin for the baseline model, and for each alternative case, (i)-(iii). Note that the observed persistence rates reported in Section 3.2 (0.77 and 0.62) are almost identical as predictions from the baseline model. The difference in the estimated persistence rate is reported for each alternative case (compared to the baseline model), and can be interpreted as the contribution of the specific type of state dependence. For reference, we have also reported the persistence rate that follows from random transitions, where the transition probabilities, $\varphi(h_t|h_{t-1})$, equals the state probabilities, $\varphi(h_t)$. In this case there are no observed or unobserved heterogeneity which creates persistence, and no preference dependence, labor market constraint dependence or switching costs.

Overall, we find that the switching cost, despite its minimalist setup, contribute most to persistence. Restricting the switching cost to zero leads to large drops in the persistence rates: from 0.76 to 0.51 and from 0.62 to 0.28 at the extensive and intensive margin respectively. Dependence via labor market opportunities also contributes substantially to explaining persistence at the extensive margin, where the rate is reduced from 0.76 to 0.53. Recall that this component of state dependence does not affect the intensive margin decisions; see Section 2. Preference dependence is on average less important. Under the restriction that past labor supply has no impact on the leisure preferences, the persistence rates are still significantly reduced both at the extensive and intensive margin by 0.15 and 0.04 respectively.[25]

A decomposition of the simulated adjustment path is provided in Panel B in Fig. 1. The short-vs-long term elasticity ratio in Table 4 also suggest that omitting switching costs contribute the most to the sluggish response path, as the elasticity ratio is increased from 0.36 to 0.53 at the extensive margin, and from 0.39 to 0.77 at the intensive margin.

4.5. Robustness checks

As discussed in section (2.4), functional form assumptions are required to identify the model. In the present subsection, we investigate the robustness of our main results by estimating different versions of the model with alternative parametric assumptions.

In Table 5 the results of our baseline model are summarized as an overall (unconditional) short term/long term elasticity of 0.39. When we omit preference dependence (the method is described in Section 4.4) the ratio is increased by 0.17 to a ratio of 0.56. If we instead omit labor market constraint dependence the ratio is increased by 0.01 to a ratio of 0.40. And if we omit switching costs the ratio is increased by 0.38 to a ratio of 0.77. This nicely summarize our main conclusion regarding the overall magnitude of sluggishness in labor market decisions, and the relative importance among different components of state dependence.

Now, to check whether these main conclusions are robust, we estimate a number of models with alternative parametric assumptions. First, we replace the box-cox utility function by a second-order polynomial utility function. Second, we apply an alternative specification of $g()$ on the basis of an a-priori distribution of job offers (not estimated within the model). Third, we estimate an alternative specification of switching costs, where only shifting between broader categories of working hours matters. Fourth, we increase the number of unobserved individual types from three (baseline model) to four. And fifth, we test a more elaborate specification of the wage rate, in which each individual has five different predicted wage rates depending on the previous period's working time choice.[26]

The results from these robustness tests are presented in Table 5. The results suggest that our main findings are quite robust across the different specifications. There is little variation in the predicted ratio between the short term and long term elasticities across the different model specifications. And the patterns on the relative contributions from different

[22] See also Haan (2010) and Gelber et al. (2020) for similar conclusions.

[23] Thoresen and Vattø (2015) follow an approach along these lines to validate the standard one-period model.

[24] Gelber et al. (2020) concludes for instance that the long-term (frictionless) estimate of earnings responses among social security recipients is nearly twice as large as the short-term impact.

[25] A complete decomposition of the persistence rate is not possible because of non-linearities; consequently, we cannot add up the contribution of each component in order to obtain the persistence rate of the baseline model.

[26] In our main analysis we use a pooled Heckman selection regression (Table A.1 in Appendix) to assign wage rates to each individual. However, one might expect that state dependence and unobserved heterogeneity is also present in wages (see e.g. Eckstein and Wolpin, 1989). In the fifth specification we therefore test a more elaborate specification of the wage rate, in which wage rates depend directly on the previous period's working time choice. Random effects are also added to the predicted individual wage rates on the basis of draws (30 draws per individual) from a normal distribution of the individual specific effects.

Table 4
Persistence rates and elasticity ratios. The relative contribution of preference dependence, labor market constraint dependence and switching costs.

		Extensive margin		Intensive margin	
		Persistence rate	Elasticity ratio	Persistence rate	Elasticity ratio
	Baseline model	0.76	0.36	0.62	0.39
(i)	Omit preference dependence	-0.15	+0.11	-0.04	+0.17
(ii)	Omit constraint dependence	-0.23	+0.14	-	-
(iii)	Omit switching costs	-0.25	+0.17	-0.34	+0.38
	Random transitions	-0.73		-0.45	

Notes: The persistence rate is defined as the share of individuals in the same working time arrangement as the previous period (see Section 3.2). Preference dependence is omitted by replacing $\hat{b_{2j}}$ by a constant average, $\bar{b_2}$. Labor market constraint dependence is omitted by replacing $\hat{\gamma_{2j}}$ by a constant average, $\bar{\gamma_2}$. Switching costs are omitted by replacing $\hat{s}$ by 0. Random transitions refers to the case when transition probabilities, $\varphi(h_t|h_{t-1})$,equals the state probabilities, $\varphi(h_t)$. In this case there are no observed or unobserved heterogeneity which creates persistence, and no preference dependence, labor market constraint dependence or switching costs.

Table 5
Robustness test of alternative parametric specifications.

	Ratio - short term/long-term elasticities			
	Full model	Omit preference dependence	Omit constraint dependence	Omit switching costs
Baseline	0.39	+0.17	+0.01	+0.38
Alt. utility function	0.39	+0.27	+0.08	+0.26
Alt. cost of switching	0.40	+0.14	+0.04	+0.22
Alt. opportunity density	0.36	+0.17	+0.02	+0.32
Alt. # unobserved types	0.38	+0.18	+0.02	+0.32
Alt. wage specification	0.36	+0.18	+0.02	+0.33

Notes: The ratios of short term/long-term elasticities are simulated by a general tax cut. Preference dependence is omitted by replacing $\hat{b_{2j}}$ by a constant average, $\bar{b_2}$. Labor market constraint dependence is omitted by replacing $\hat{\gamma_{2j}}$ by a constant average, $\bar{\gamma_2}$. Switching costs are omitted by replacing $\hat{s}$ by 0.Alt. utility function refers to that the box-cox utility function is replaced by a second-order polynomial utility function. Alt. cost of switching refers to a specification where only shifting between broader categories of working hours matters. Alt. opportunity density refers to an alternative specification of g() on the basis of an a-priori distribution of job offers (not estimated within the model). Alt. # of unobserved types refers to a model with 4 unobserved types instead of three. Alt. wage specification refers to a more elaborate specification of the wage rate, in which each individual has five different predicted wage rates depending on the previous period' s working time choice.

components are also similar. All specifications (slight exception in the first alternative specification) suggest that switching costs (status quo bias) is the most important contributor to sluggish responsiveness.

4.6. Policy relevance

In this section we summarize the main policy implications of our results. First, we describe what our results imply about the time path of responses, second, we discuss the policy relevance of decomposing results into different mechanisms of state dependence in labor supply decisions.

First, our results suggest that the expected effects of policy changes takes time to be fully realized. This implies that policymakers should be made aware of not only the potential effects, but also how fast these effects can be expected to be reached. One example is the revenue estimation of proposed tax changes. Politicians that favor tax cuts often refer to the positive effects on economic activity and argue that their proposed tax cut is not as costly as when accounting for behavioral responses. However, as our paper suggested, static behavioral models made available for policy makers may overstate the short-term behavioral responses as labor supply outcomes are predicted under the assumption that workers can adjust their behavior freely irrespective of their status quo choices. Similarly, quasi-experimental evidence is likely to understate the long-term effect of policy changes. Thus, our results suggest that economists offering guidance to policy makers should be clearer on whether their predictions are likely to reflect short-term or long-term effects.

Second, what is the policy implication of decomposing results into different categories of state dependence? First of all, it can be noted that in the public economics literature there has been much attention on frictions which prevents people from optimizing. But sluggish responses can also be caused by preference dependence such as innate sluggishness in people's preferences or habit persistence which from the individual's perspective is optimal. We therefore argue that while preference dependence can be seen as optimal for the individual (in a welfare sense), the sluggishness induced by labor market constraints and status quo bias (switching costs) hinder individuals to choose their optimal labor supply choices, which leads to a welfare loss for the society. It is still important to try to distinguish the two latter categories as they call for different labor market policies to reduce welfare losses: In the case of strong state dependence via labor market constraints, one may want to consider demand side policies, which reduces the cost of employment by introducing wage subsidies or reducing social security contributions. In contrast, if status quo bias (switching cost) plays an important role, as we find in our paper, one might want to focus more on supply side policies that reduce the switching cost, such as providing more information to potential job seekers (Altmann et al., 2018). Also, one might want to resort to behavioral studies and consider possible “nudges” to prevent “irrational” status quo bias and develop more cost-effective policies, see Belot et al. (2019) and Babcock et al. (2012).

To summarize, our results suggest that sluggishness in responses is to a large extent caused by status quo bias (switching costs) which can potentially be altered by policy, and to a smaller degree to innate preference dependence caused by “optimal” sluggishness in people' s responses. We acknowledge that we rely on functional form assumptions to distinguish the different components, nevertheless, as shown in Section 4.5, our main conclusion seems to be rather robust to alternative functional form specifications.

5. Conclusion

The aim of this paper is to develop a model that explicitly takes account of state dependence in labor supply decisions in order to better understand the labor supply behavior in response to policy changes and simulate the path of labor supply adjustments. We consider three possible channels in which past labor supply behavior impacts current decisions: via preferences, labor market constraints and (fixed) switching costs, while controlling for observed and unobserved individual heterogeneity. All three elements cause gradual responses to policy changes until a new optimum is reached.

The model is estimated for Norwegian women and it reproduces the empirical patterns observed in the data. The estimated model is used to map out the adjustment path to a hypothetical permanent tax cut. We find that the estimated first-year responses are brought down to one-third of the long-term effect, which is reached after about five years. In the long-term (when the new optimum is reached), simulated labor supply responses are close to predictions from a standard one-period model where workers are assumed to adjust their behavior freely irrespective of their initial choices.

In addition to contributing to the literature regarding the time path of labor supply adjustments, our model represents the first attempt in the literature to distinguish between the different mechanisms of state dependence in a structural framework. We argue that modeling state dependence in this matter both improves the predictive power of our model, but also offers a unique opportunity to study the relative contributions of the different mechanism of sluggish responses to policy changes.

Our results suggest that both preference dependence, labor market constraint dependence and status quo bias (switching costs) matter for the predicted sluggish labor market responses. We find that status quo bias (switching cost) is the most important component of sluggish responses. We argue that each component of state dependence calls for different policy implications: Whereas preference dependence can be considered optimal from the individual' s perspective, labor market constraint dependence and status quo bias represent frictions which prevents people from choosing their optimal labor market outcomes. Labor market constraint dependence is likely to be mitigated by labor demand policies which reduces the cost of employment, whereas status quo bias can be mitigated by improving information or by considering possible “nudges” to push people over to their optimal choices.

We acknowledge that there are two clear limitations to our approach. First of all, the different components of state dependence are not non-parametrically identified. This means that we need to rely on functional form assumptions to disentangle the different components of state dependence. Nevertheless, we perform the same analysis using various functional form assumptions and the results suggest that our findings are robust. The second limitation of our study is that our study abstracts from forward looking behavior. Our model framework assumes that individuals maximize utility on a year-to-year basis, such that agents are short-sighted and do not consider that choices today affect future frictions. A completely forward-looking model would predict anticipatory adjustment and intertemporal substitution in addition to the sluggish responses we focus on in the present paper. Thus, a challenging question for future research on this topic is whether the assumption of myopic agents can be relaxed to incorporate all these dynamic effects into the path of labor supply adjustment.

To conclude, this paper adds to an emerging literature which suggests that individuals respond gradually rather than immediately to policy changes due to frictions or state dependence. We contribute to the existing literature by explicitly modeling different components of state dependence in a structural model framework. This offers a particular opportunity to study the relative contributions of the different mechanism of sluggish responses, as well as improving predictions for practical tax policy.

Appendix

Imputation of missing working time observations

Missing working time observations (about 30 percent) are imputed from observed information on annual labor income combined with predicted monthly wage income for a full-time job. Monthly wage income is predicted from a set of individual characteristics (experience, field and level of education, National background and county). Cut-offs for each working time choice are calibrated by adjusting the simulated to

Table A.1
Wage regression, Heckman 2-Stage.

	Log wage		Participation	
	Coefficient	Std. error	Coefficient	Std. error
Constant	4.6847***	(0.0069)	1.5124***	(0.0674)
Experience	0.0149***	(0.0003)	0.0109**	(0.0042)
Experience squared	-0.0002***	(0.0000)	-0.0009***	(0.0001)
Low education	-0.0627***	(0.0015)	-0.2360***	(0.0164)
Higher education	0.2573***	(0.0016)	0.3004***	(0.0214)
Non-western immigrants	-0.1183***	(0.0037)	-0.9665***	(0.0231)
Residence in metropolitan area	0.0764***	(0.0013)	-0.0826***	(0.0141)
Educational category				
General	0.0177***	(0.0050)	0.6694***	(0.0306)
Humanities, arts	-0.0311***	(0.0055)	0.5069***	(0.0394)
Education	-0.0450***	(0.0054)	1.0106***	(0.0423)
Social studies, law	0.1097***	(0.0062)	1.0089***	(0.0637)
Business	0.0696***	(0.0052)	0.8846***	(0.0327)
Technology	0.1098***	(0.0054)	0.8938***	(0.0386)
Health	-0.0418***	(0.0052)	1.0955***	(0.0329)
Primary industries	0.0445***	(0.0091)	0.3778***	(0.0736)
Services	0.0179**	(0.0065)	0.6057***	(0.0494)
Exclusion restrictions				
No. of children under age 3			-0.1210***	(0.0201)
No. of children under age 6			-0.0540**	(0.0171)
No. of children under age 12			-0.3400***	(0.0095)
Net wealth (NOK 10,000)			-0.0009***	(0.0001)
Partner's net income (NOK 10,000)			-0.0032***	(0.0001)
Mills Lambda	0.0318***	(0.0056)		
Number of observations	121,408		167,272	
Number of individuals	20,466		23,896	

Notes: The wage regression is used to assign each individual a wage rate before estimating the baseline labor supply model. The dependent variable is the hourly wage rate calculated by dividing contractual monthly pay by contractual monthly working hours. The regression includes year fixed effects, and allows for participation selection effects. The educational category "unknown" serves as the reference category. Wages, net wealth and partner's net income are measured in current NOK (NOK 1≈USD 0.17≈EUR 0.13). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.2
Predicted annual transition probabilities in the period 2003 - 2009. All individuals.

Year t	Year t + 1 None	Short p-t	Long p-t	Full-time	Overtime
None	**83.5%**	12.1%	4.1%	0.3%	0.0%
Short p-t	2.9%	**75.9%**	15.4%	5.5%	0.2%
Long p-t	0.1%	8.8%	**75.1%**	14.0%	2.1%
Full-time	0.0%	2.6%	7.9%	**81.3%**	8.1%
Overtime	0.0%	4.6%	11.1%	23.5%	**60.7%**
Persistence rate	**0.76**	**0.61**	**0.61**	**0.68**	**0.42**

Notes: The predicted transition probabilities are based on an average of the predicted annual transition for the sample of married women over the period 2003 - 2009. The working time categories are based on working hours per week: 0 (“None”), 1–19 (“Short part-time”), 20–34 (“Long part-time”), 35–40 (“Full-time”) and 41 + (“Overtime”).

Table A.3
Predicted and observed one-year transition probabilities (2004–2005). Low wage individuals.

	Predicted				
	Year t + 1				
Year t	None	Short p-t	Long p-t	Full-time	Overtime
None	**85.7%**	10.8%	3.2%	0.2%	0.0%
Short p-t	3.8%	**78.3%**	13.3%	4.4%	0.2%
Long p-t	0.2%	10.4%	**75.2%**	12.5%	1.8%
Full-time	0.1%	4.0%	8.6%	**79.5%**	7.8%
Overtime	0.2%	4.9%	11.1%	22.6%	**61.2%**
	Observed				
	Year t+1				
Year t	None	Short p-t	Long p-t	Full-time	Overtime
None	**84.5%**	13.6%	1.0%	0.8%	0.2%
Short p-t	3.8%	**79.6%**	12.7%	3.2%	0.7%
Long p-t	0.1%	10.7%	**77.6%**	9.5%	2.0%
Full-time	0.0%	3.4%	11.8%	**77.1%**	7.8%
Overtime	0.0%	3.4%	7.4%	34.2%	**55.0%**

Notes: The predicted transition probabilities are based on an average of the predicted annual transition for the sample of married women over the period 2003 - 2009. The working time categories are based on working hours per week: 0 (“None”), 1–19 (“Short part-time”), 20–34 (“Long part-time”), 35–40 (“Full-time”) and 41 + (“Overtime”). Individuals are categorized into low wage (below median) and high wage (above median) individuals according to their predicted hourly wage rate in year 2003.

Table A.4
Predicted and observed one-year transition probabilities (2004–2005). High wage individuals.

	Predicted				
	Year t + 1				
Year t	None	Short p-t	Long p-t	Full-time	Overtime
None	**80.6%**	13.3%	5.7%	0.4%	0.0%
Short p-t	2.4%	**72.9%**	17.4%	7.0%	0.3%
Long p-t	0.1%	7.4%	**74.7%**	15.3%	2.4%
Full-time	0.0%	2.2%	7.2%	**82.1%**	8.4%
Overtime	0.0%	3.1%	10.2%	23.9%	**62.8%**
	Observed				
	Year t+1				
Year t	None	Short p-t	Long p-t	Full-time	Overtime
None	**84.2%**	13.2%	2.1%	0.5%	0.0%
Short p-t	2.8%	**69.7%**	21.1%	5.4%	1.1%
Long p-t	0.0%	7.6%	**74.5%**	15.4%	2.5%
Full-time	0.0%	1.7%	10.5%	**81.0%**	6.8%
Overtime	0.0%	1.5%	7.4%	41.7%	**49.4%**

Notes: The predicted transition probabilities are based on an average of the predicted annual transition for the sample of married women over the period 2003 - 2009. The working time categories are based on working hours per week: 0 (“None”), 1–19 (“Short part-time”), 20–34 (“Long part-time”), 35–40 (“Full-time”) and 41 + (“Overtime”). Individuals are categorized into low wage (below median) and high wage (above median) individuals according to their predicted hourly wage rate in year 2003.

Table A.5
Out of sample fit. Predicted and observed five-years transition probabilities. 2004–2009.

	Predicted				
	Year t + 1				
Year t	None	Short p-t	Long p-t	Full-time	Overtime
None	**55.9%**	24.7%	12.8%	5.7%	1.0%
Short p-t	3.9%	**47.0%**	35.5%	12.8%	0.8%
Long p-t	0.5%	14.5%	**52.0%**	29.8%	3.3%
Full-time	0.3%	5.5%	17.2%	**66.6%**	10.4%
Overtime	0.5%	3.3%	12.0%	42.4%	**42.0%**
	Observed				
	Year t+1				
Year t	None	Short p-t	Long p-t	Full-time	Overtime
None	**54.2%**	29.1%	9.7%	5.2%	1.8%
Short p-t	4.1%	**48.1%**	34.8%	10.5%	2.5%
Long p-t	0.7%	9.5%	**56.9%**	27.2%	5.7%
Full-time	0.5%	3.7%	13.4%	**72.8%**	9.5%
Overtime	0.5%	4.2%	14.1%	43.6%	**37.6%**

Notes: The estimated model (estimated on 10% sample) is used to compare simulated (“predicted”) and observed outcome in a separate 3% sample over the period 2004 - 2009. The working time categories are based on working hours per week: 0 (“None”), 1–19 (“Short part-time”), 20–34 (“Long part-time”), 35–40 (“Full-time”) and 41 + (“Overtime”).

Table A.6
Out of sample fit. Predicted and observed five-years transition probabilities of a semi-parametric model. 2004–2009.

	Predicted				
	Year t + 1				
Year t	None	Short p-t	Long p-t	Full-time	Overtime
None	**53.3%**	27.2%	11.5%	5.8%	2.2%
Short p-t	4.2%	**48.8%**	31.0%	13.6%	2.4%
Long p-t	0.7%	10.3%	**58.2%**	23.9%	6.9%
Full-time	0.3%	4.8%	14.1%	**70.9%**	9.8%
Overtime	0.5%	3.8%	18.4%	40.4%	**36.9%**
	Observed				
	Year t+1				
Year t	None	Short p-t	Long p-t	Full-time	Overtime
None	**54.2%**	29.1%	9.7%	5.2%	1.8%
Short p-t	4.1%	**48.1%**	34.8%	10.5%	2.5%
Long p-t	0.7%	9.5%	**56.9%**	27.2%	5.7%
Full-time	0.5%	3.7%	13.4%	**72.8%**	9.5%
Overtime	0.5%	4.2%	14.1%	43.6%	**37.6%**

Notes: The estimated Semi-parametric model (estimated on 10% sample) is used to compare simulated (“predicted”) and observed outcome in a separate 3% sample over the period 2004 - 2009. The Semi-parametric model is estimated by estimating a parameter for each possible one-year transition. The working time categories are based on working hours per week: 0 (“None”), 1–19 (“Short part-time”), 20–34 (“Long part-time”), 35–40 (“Full-time”) and 41 + (“Overtime”).

Table A.7
Persistence rates and elasticity ratios. Out-of-sample comparison of our model and a semi-parametric model.

	Extensive margin		Intensive margin	
	Persistence rate	Elasticity ratio	Persistence rate	Elasticity ratio
Observed	0.78		0.62	
Our baseline model	0.77	0.36	0.62	0.39
Semi-parametric model	0.77	0.27	0.61	0.31

Notes: The persistence rate is here defined as the share of individuals (out-of-sample) in the same working time arrangement in 2004 and 2009.

the actual distribution of working time for individuals observed in the wage statistics sample. We use the following cut-offs: short part-time: 0.3–7.7 times predicted monthly wage, long part-time: 7.7–10.95, full-time: 10.95-15.35 and overtime: >15.35. Non-participants are defined as earning less than NOK 5000 (about USD 850 or EUR 530) annually or less than 0.3 times predicted full-time monthly earnings.

Assigning wage rates and wage income for each working time option

Table A.1 presents the wage regression, which is used to assign each individual a wage rate. The dependent variable in the regression is the hourly wage rate which is calculated by dividing contractual monthly pay by contractual monthly working hours. We use a (pooled) Mincer wage regression with year fixed effects, allowing for participation selection effects (Heckman, 1979).[27] We assign the predicted hourly wage rate of the Heckman regression to each individual. Wage income is constructed by multiplying the individual wage rate and the median (annual) working time for each option.

References

Altmann, S., Falk, A., Jäger, S., Zimmermann, F., 2018. Learning about job search: a field experiment with job seekers in germany. J. Public Econ. 164, 33–49.

Arulampalam, W., Booth, A.L., Taylor, M.P., 2000. Unemployment persistence. Oxf. Econ. Pap. 52 (1), 24–50.

Babcock, L., Congdon, W.J., Katz, L.F., Mullainathan, S., 2012. Notes on behavioral economics and labor market policy. IZA Journal of Labor Policy 1 (1), 1–14.

Bækgaard, H., 2014. The differences-in-differences approach with overlapping differences-experimental verification of estimation bias. DREAM Working Paper 2014:3.

Belot, M., Kircher, P., Muller, P., 2019. Providing advice to jobseekers at low cost: an experimental study on online advice. Rev. Econ. Stud. 86 (4), 1411–1447.

Bloemen, H.G., 2008. Job search, hours restrictions, and desired hours of work. J. Labor Econ. 26 (1), 137–179.

Chetty, R., 2012. Bounds on elasticities with optimization frictions: a synthesis of micro and macro evidence on labor supply. Econometrica 80 (3), 969–1018.

Chetty, R., Friedman, J.N., Olsen, T., Pistaferri, L., 2011. Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: evidence from danish tax records. Q. J. Econ. 126 (2), 749–804.

Creedy, J., Kalb, G., 2005. Discrete hours labour supply modelling: specification, estimation and simulation. J. Econ. Surv. 19 (5), 697–734.

Dagsvik, J., Jia, Z., Kornstad, T., Thoresen, T.O., 2014. Theoretical and practical arguments for modeling labor supply as a choice among latent jobs. J. Econ. Surv. 28 (1), 134–151.

Dagsvik, J.K., Jia, Z., 2016. Labor supply as a choice among latent jobs: unobserved heterogeneity and identification. Journal of Applied Econometrics 31 (3), 487–506.

Dobson, J., Henthorne, K., Lynas, Z., 2000. Pupil mobility in schools: Final report. Migration Research Unit, UCL.

Dubé, J.-P., Hitsch, G.J., Rossi, P.E., 2010. State dependence and alternative explanations for consumer inertia. Rand J. Econ. 41 (3), 417–445.

Dunn, L.F., 1996. Loss aversion and adaptation in the labor market: empirical indifference functions and labor supply. Rev. Econ. Stat. 441–450.

Eckstein, Z., Wolpin, K.I., 1989. Dynamic labour force participation of married women and endogenous work experience. Rev. Econ Stud. 56 (3), 375–390.

Eriksson, S., Rooth, D.-O., 2014. Do employers use unemployment as a sorting criterion when hiring? evidence from a field experiment. American Economic Review 104, 1014–1039. doi:10.1257/aer.104.3.1014.

Euwals, R., Van Soest, A., 1999. Desired and actual labour supply of unmarried men and women in the netherlands. Labour Econ. 6 (1), 95–118.

Furnham, A., Boo, H.C., 2011. A literature review of the anchoring effect. J. Socio. Econ. 40 (1), 35–42.

Gelber, A.M., Jones, D., Sacks, D.W., 2020. Estimating adjustment frictions using nonlinear budget sets: method and evidence from the earnings test. American Economic Journal: Applied Economics 12 (1), 1–31.

Giertz, S.H., 2010. The elasticity of taxable income during the 1990s: new estimates and sensitivity analyses. South. Econ. J. 77 (2), 406–433.

Haan, P., 2010. A multi-state model of state dependence in labor supply: intertemporal labor supply effects of a shift from joint to individual taxation. Labour Econ. 17 (2), 323–335.

Haan, P., Kemptner, D., Uhlendorff, A., 2015. Bayesian procedures as a numerical tool for the estimation of an intertemporal discrete choice model. Empir. Econ. 49 (3), 1123–1141.

Haan, P., Uhlendorff, A., 2013. Intertemporal labor supply and involuntary unemployment. Empir. Econ. 44 (2), 661–683.

Heckman, J.J., 1979. Sample selection bias as a specification error. Econometrica 47 (1), 153–162.

Heckman, J.J., 1981. Statistical Models for Discrete Panel Data Models. In: Manski, C.F., McFadden, D. (Eds.), Structural Analysis of Discrete Data with Econometric Applications. MIT Press Cambridge, MA, pp. 114–175.

Heckman, J.J., 1993. What has been learned about labor supply in the past twenty years? Am. Econ. Rev. 83 (2), 116–121.

Heckman, J.J., Borjas, G.J., 1980. Does unemployment cause future unemployment? definitions, questions and answers from a continuous time model of heterogeneity and state dependence. Economica 47 (187), 247–283.

Hyslop, D.R., 1999. State dependence, serial correlation and heterogeneity in intertemporal labor force participation of married women. Econometrica 67 (6), 1255–1294.

Jongen, E.L., Stoel, M., 2019. The elasticity of taxable labour income in the netherlands. Economist (Leiden) 167 (4), 359–386.

Kleven, H.J., Schultz, E.A., 2014. Estimating taxable income responses using danish tax reforms. American Economic Journal: Economic Policy 6 (4), 271–301.

Kleven, H.J., Waseem, M., 2013. Using notches to uncover optimization frictions and structural elasticities: theory and evidence from pakistan. Q. J. Econ. 128 (2), 669–723.

Kubin, I., Prinz, A., 2002. Labour supply with habit formation. Econ. Lett. 75 (1), 75–79.

Lockwood, B., 1991. Information externalities in the labour market and the duration of unemployment. Rev. Econ. Stud. 58 (4), 733–753.

Low, H., Meghir, C., 2017. The use of structural models in econometrics. Journal of Economic Perspectives 31 (2), 33–58.

Matejka, F., McKay, A., 2014. Rational inattention to discrete choices: anew foundation for the multinomial logit model. Am. Econ. Rev. 105 (1), 272–298.

Neisser, C., 2017. The elasticity of taxable income: a meta-regression analysis. ZEW Discussion Papers 17.

Neuman, T., Neuman, E., Neuman, S., 2010. Explorations of the effect of experience on preferences for a health-care service. J. Socio. Econ. 39 (3), 407–419. https://doi.org/10.1016/j.socec.2010.02.005.

Parady, G., Ory, D., Walker, J., 2021. The overreliance on statistical goodness-of-fit and under-reliance on model validation in discrete choice models. Journal of Choice Modelling 38. doi:10.1016/j.jocm.2020.100257.

Pissarides, C.A., 1992. Loss of skill during unemployment and the persistence of employment shocks. Q. J. Econ. 107 (4), 1371–1391.

Pronzato, C.D., 2015. Comparing quasi-experimental designs and structural models for policy evaluation: the case of a reform of lone parental welfare. CESifo Econ. Stud. 61 (1), 95–122.

Prowse, V., 2012. Modeling employment dynamics with state dependence and unobserved heterogeneity. Journal of Business & Economic Statistics 30 (3), 411–431.

Samuelson, W., Zeckhauser, R., 1988. Status quo bias in decision making. J. Risk Uncertain. 1 (1), 7–59.

San Miguel, F., Ryan, M., Scott, A., 2002. Are preferences stable? the case of health care. Journal of Economic Behavior & Organization 48 (1), 1–14.

Schorfheide, F., Wolpin, K.I., 2016. To hold out or not to hold out. Research in Economics 70, 332–345. doi:10.1016/j.rie.2016.02.001.

Seetharaman, P., 2004. Modeling multiple sources of state dependence in random utility models: a distributed lag approach. Marketing Science 23 (2), 263–271.

Skrondal, A., Rabe-Hesketh, S., 2004. Generalized latent variable modeling: multilevel, longitudinal, and structural equation models. Chapman & Hall/CRC.

Spence, M., 1973. Job market signaling. Quarterly Journal of Economics 87, 355–374.

Steiner, J., Stewart, C., Matějka, F., 2017. Rational inattention dynamics: inertia and delay in decision-making. Econometrica 85 (2), 521–553.

Thoresen, T.O., Vattø, T.E., 2015. Validation of the discrete choice labor supply model by methods of the new tax responsiveness literature. Labour Econ. 37, 38–53.

Van Soest, A., 1995. Structural models of family labor supply: adiscrete choice approach. Journal of Human Resources 30 (1), 63–88.

Van Soest, A., Das, M., Gong, X., 2002. A structural labour supply model with flexible preferences. J. Econom. 107, 345–374. doi:10.1016/S0304-4076(01)00128-2.

Vattø, T.E., 2020. Estimating the elasticity of taxable income when earnings responses are sluggish. FinanzArchiv/ Public Finance Analysis 76 (4), 329–369.

Vishwanath, T., 1989. Job search, stigma effect, and escape rate from unemployment. J. Labor Econ. 7 (4), 487–502.

Woittiez, I., Kapteyn, A., 1998. Social interactions and habit formation in a model of female labour supply. J. Public Econ. 70 (2), 185–205.

Wooldridge, J.M., 2005. Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. Journal of Applied Econometrics 20 (1), 39–54.

[27] The Mills lambda value is significantly positive, suggesting that unobservables are positively related to both labor market participation and wages.