Online Appendix to "The Optimal Timing of Unemployment Benefits: Theory and Evidence from Sweden"

Jonas Kolsrud, Uppsala
Camille Landais, London School of Economics
Peter Nilsson, IIES
Johannes Spinnewijn, London School of Economics

FOR ONLINE PUBLICATION - Appendix A: Technical Appendix

This Appendix provides the technical details of our model setup and the proofs of the Propositions. We also consider extensions of the baseline model to gauge the robustness of our characterization of the optimal benefit profile. We finally compare the evolution of consumption smoothing gains and moral hazard in stationary and non-stationary environments.

A.1 Setup

We closely follow the setup in Chetty [2006], but allow for heterogeneous agents and non-stationarities. Let $\omega_{i,t}$ denote a vector of state variables that contain all relevant information up to time t in determining an agent i's employment status and behavior at time t. Let $F_{i,t}(\omega_{i,t})$ denote the unconditional distribution of $\omega_{i,t}$ given information available at time 0. We assume that $F_{i,t}$ is a smooth function and let Ω denote the maximal support of $F_{i,t}$ for $\forall i, \forall t$. In our stylized model, the vector of state variables $\omega_{i,t}$ includes only the asset level, time and the employment status.

In each period t, an agent decides how much to consume from her income and assets. In our stylized model, an agent earns $w - \tau$ when employed and receives b when unemployed. The law of motion of assets in the employment and unemployment state are respectively,

$$a_{i,t+1} = ra_{i,t} + w - \tau - c^e_{i,t} \tag{1}$$

$$a_{i,t+1} = ra_{i,t} + b_t - c^u_{i,t}, \tag{2}$$

but are constrained to be above $a_{i,t+1} \geq \bar{a}_i$ for each agent i and any time t. We denote the Lagrange multipliers on these constraints by $\mu^e_{i,t}(\omega_{i,t}), \mu^u_{i,t}(\omega_{i,t})$ and $\mu^a_{i,t}(\omega_{i,t})$ respectively.

Let $\theta_{i,t}(\omega_{i,t})$ denote an agent's employment status at time t in state $\omega_{i,t}$. If $\theta = 1$, the agent is employed, and if $\theta = 0$, the agent is unemployed. In each period t, an unemployed agent chooses a level of search effort $s_{i,t}$ as well. This search effort level determines the probability to leave unemployment for employment in that period. This mapping may be agent-specific and change depending on the length of the unemployment spell.

Each agent i chooses a program (s_i, c_i^u, c_i^e) with

$$s_i = \{s_{i,t}(\omega_{i,t})\}_{t\in\{1,2,..T\},\omega_{i,t}\in\Omega,\theta(\omega_{i,t})=0},$$
$$c_i^u = \{c_{i,t}^u(\omega_{i,t})\}_{t\in\{1,2,..T\},\omega_{i,t}\in\Omega,\theta(\omega_{i,t})=0},$$
$$c_i^e = \{c_{i,t}^e(\omega_{i,t})\}_{t\in\{1,2,..T\},\omega_{i,t}\in\Omega,\theta(\omega_{i,t})=1},$$

to solve

$$\begin{aligned}V_i(P) = \max \Sigma_{t=1}^T \beta^{t-1} \int \{ & v_i^u\left(c_{i,t}^u(\omega_{i,t}), s_{i,t}(\omega_{i,t})\right)[1-\theta_{i,t}(\omega_{i,t})] + v_i^e\left(c_{i,t}^e(\omega_{i,t})\right)\theta_{i,t}(\omega_{i,t})\} dF_{i,t}(\omega_{i,t}) \\
&+ \Sigma_{t=1}^T \beta^{t-1} \int \mu_{i,t}^u(\omega_{i,t})\left[ra_{i,t}(\tilde{\omega}_{i,t-1}) + b_t - c_{i,t}^u(\omega_{i,t}) - a_{i,t+1}(\omega_{i,t})\right][1-\theta_{i,t}(\omega_{i,t})]\, dF_{i,t}(\omega_{i,t}) \\
&+ \Sigma_{t=1}^T \beta^{t-1} \int \mu_{i,t}^e(\omega_{i,t})\left[ra_{i,t}(\tilde{\omega}_{i,t-1}) + w - \tau - c_{i,t}^e(\omega_{i,t}) - a_{i,t+1}(\omega_{i,t})\right]\theta_{i,t}(\omega_{i,t})\, dF_{i,t}(\omega_{i,t}) \\
&+ \Sigma_{t=1}^T \beta^{t-1} \int \mu_{i,t}^a(\omega_{i,t})\left[\bar{a}_i - a_{i,t+1}(\omega_{i,t})\right] dF_{i,t}(\omega_{i,t}),\end{aligned}$$

where we use the short-hand notation $\tilde{\omega}_{i,t-1}$ to denote the vector of state variables at time $t-1$ that preceded the vector of state variables $\omega_{i,t}$ at time t. Following Chetty [2006], we assume that lifetime utility is smooth, increasing and strictly quasi-concave in (c_i^u, c_i^e, s_i) and that the value function $V_i(P)$ is differentiable such that the Envelope Theorem applies. This implies that

$$\begin{aligned}\frac{\partial V_i(P)}{\partial b_t} &= \beta^{t-1}\int \mu_{i,t}^u(\omega_{i,t})[1-\theta_{i,t}(\omega_{i,t})]\, dF_{i,t}(\omega_{i,t}) \\
&= \beta^{t-1}\int \frac{\partial v_i^u\left(c_{i,t}^u(\omega_{i,t}), s_{i,t}(\omega_{i,t})\right)}{\partial c_{i,t}^u}[1-\theta_{i,t}(\omega_{i,t})]\, dF_{i,t}(\omega_{i,t}).\end{aligned}$$

The second equality uses the optimality of the consumption choice $c_{i,t}^u(\omega_{i,t})$, which does not depend on the borrowing constraint being binding or not.

In our stylized model, the agent starts unemployed and remains employed until T once she finds a job. The agent's exit rate out of unemployment at time t only depends on her search effort at time t. The (unconditional) probability to be unemployed at time $t+1$ therefore simplifies to

$$\Pr(\theta_{i,t+1}=0) = \int (1 - h_{i,t}(s_{i,t}(\omega_{i,t})))[1-\theta_{i,t}(\omega_{i,t})]\, dF_{i,t}(\omega_{i,t}).$$

This simplifying assumption makes that on the optimal path an agent's unemployment consumption $c_{i,t}^u(\omega_{i,t})$ only varies with time t, which coincides with the number of periods she is currently unemployed. The agent's employment consumption $c_{i,t}^e(\omega_{i,t})$, however, depends on both time t and the number of periods she has been unemployed.

We now turn to the policy. We can express the present value of the government's budget as

$$G(P) = \Sigma_{t=1}^T [1+r]^{-(t-1)} \int\int \{-b_t[1-\theta_{i,t}(\omega_{i,t})] + \tau\theta_{i,t}(\omega_{i,t})\} dF_{i,t}(\omega_{i,t})\, di,$$

which simplifies to (1) when $r = 0$.

The government solves

$$\max \int V_i(P)\, di + \lambda\left[G(P) - \bar{G}\right],$$

where λ is the Lagrange multiplier on the government's budget constraint and $\bar{G}$ is an exogenous revenue constraint. Our characterization is based on local policy changes and thus only allows for local tests and recommendations. For the local recommendations to translate globally, we would need the program to be strictly concave in P.[1] To provide

[1] Chetty (2006) provides regularity conditions such that the government's problem is strictly concave in case of

tractable expressions of the local welfare implications we assume that the social welfare function is differentiable.

A.2 Dynamic Unemployment Policy

A.2.1 Proof of Proposition 1

The welfare impact of a change in benefit level b_t of policy P equals

$$\frac{\partial W(P)}{\partial b_t} = \int \frac{\partial V_i(P)}{\partial b_t} di + \lambda \frac{\partial G(P)}{\partial b_t},$$

where, using $S_t^r \equiv S_t / [1+r]^{(t-1)}$ and $\varepsilon_{t',t}^r = \frac{\partial S_{t'}^r}{\partial b_t}\frac{b_t}{S_{t'}^r}$,

$$\frac{\partial G(P)}{\partial b_t} = -S_t^r - \Sigma_{t'=1}^T (b_{t'} + \tau) \frac{\partial S_{t'}^r}{\partial b_t} = -S_t^r \times \left[1 + \Sigma_{t'=1}^T \frac{S_{t'}^r (b_{t'} + \tau)}{S_t^r b_t} \varepsilon_{t',t}^r\right],$$

which simplifies to expression (5) for $r = 0$, and

$$\begin{aligned}\int \frac{\partial V_i(P)}{\partial b_t} di &= \int \int \beta^{t-1} \frac{\partial v_i^u \left(c_{i,t}^u(\omega_{i,t}), s_{i,t}(\omega_{i,t})\right)}{\partial c_{i,t}^u} [1 - \theta_{i,t}(\omega_{i,t})] dF_{i,t}(\omega_{i,t})\} di \\ &= \beta^{t-1} S_t E\left(\frac{\partial v_i^u \left(c_{i,t}^u(\omega_{i,t}), s_{i,t}(\omega_{i,t})\right)}{\partial c_{i,t}^u} | t, \theta_{i,t}(\omega_{i,t}) = 0\right).\end{aligned}$$

This expression simplifies to (7) for $\beta = 1 + r = 1$. The expectation operator $E_t^u(\cdot)$ thus *averages over all potential states in which the agent is unemployed at time t. In our stylized setup (which assumes that the agent starts unemployed and remains employed once she finds a job), the agent's unemployment consumption $c_{i,t}^u(\omega_{i,t})$ only depends on the length of the ongoing unemployment spell. The weight of agent i's marginal utility in calculating the average marginal utility among the unemployed at time t is scaled by $S_{i,t}/S_t$.*

Combining the two expressions, we find

$$\frac{\partial W(P)}{\partial b_t} = 0 \Leftrightarrow \frac{\int \frac{\partial V_i(P)}{\partial b_t} di - \lambda}{\lambda} = \Sigma_{t'=1}^T \frac{S_{t'}^r (b_{t'} + \tau)}{S_t^r b_t} \varepsilon_{t',t}^r.$$

In the same way, we find

$$\frac{\partial G(P)}{\partial \tau} = \left[\Sigma_{t=1}^T (1 - S_t) / [1+r]^{(t-1)}\right] \times \left[1 + \Sigma_{t'=1}^T \frac{S_{t'}^r (b_{t'} + \tau)}{\Sigma_{t=1}^T \{(1 - S_t) / [1+r]^{(t-1)}\} \tau} \varepsilon_{t',\tau}^r\right],$$

$$\int \frac{\partial V_i(P)}{\partial \tau} di = \Sigma_{t=1}^T \beta^{t-1} (1 - S_t) E\left(\frac{\partial v_i^e \left(c_{i,t}^e(\omega_{i,t})\right)}{\partial c_{i,t}^e} | t, \theta_{i,t}(\omega_{i,t}) = 1\right),$$

and, hence,

$$\frac{\partial W(P)}{\partial \tau} = 0 \Leftrightarrow \frac{\lambda - \int \frac{\partial V_i(P)}{\partial \tau} di}{\lambda} = \Sigma_{t'=1}^T \frac{S_{t'}^r (b_{t'} + \tau)}{\Sigma_{t=1}^T \{(1 - S_t) / [1+r]^{(t-1)}\} \tau} \varepsilon_{t',\tau}^r,$$

which simplifies to expression (9) for $\beta = 1 + r = 1$, with $(T - D)$ equal to the expected time spent employed $\Sigma_{t=1}^T (1 - S_t)$. *The expectation operator $E^e(\cdot)$* in *(9) is over all employment states* and periods t. *Compared to consumption during unemployment, employment consumption $c_{i,t}^e(\omega_{i,t})$ depends on the unemployment history and not just on time t. Hence, we need to calculate the average marginal utility when employed at time t for any agent i and scale the weight in calculating the average marginal utility among the employed at time t by $(1 - S_{i,t}) / (1 - S_t)$. We then average over all periods t using weights $(1 - S_t) / (T - D)$.*

---

flat unemployment policies (i.e., $b_k = \bar{b}$).

The $n+1$ first-order conditions stated in the Proposition, jointly with the budget constraint, are necessary considtions for an interior, optimal policy.■

A.2.2 Robustness of Characterization

We briefly show how the optimal tax formulae continue to apply in a model with multiple unemployment spells where an agent i's layoff probability $l_i(e_{i,t})$ at time t depends on her effort on the job $e_{i,t}$. We still assume that $\omega_{i,t}$ contains all relevant information up to time t in determining an agent i's employment status and behavior at time t. Let $\theta_{i,t}(\omega_{i,t})$ still denote an agent's employment status at time t in state $\omega_{i,t}$. If $\theta=1$, the agent is employed, and if $\theta=0$, the agent is unemployed.

From the consumption smoothing perspective, the agent's marginal utility when employed can now depend on the effort on the job, $\partial v_i^e\left(c_{i,t}^e(\omega_{i,t}), e_{i,t}(\omega_{i,t})\right)/\partial c_{i,t}^e$. From the moral hazard perspective, the (unconditional) probability to be unemployed now equals

$$Pr\left(\theta_{i,t+1}=0\right)=\int\left\{\left(1-h_i\left(s_{i,t}\left(\omega_{i,t}\right),\omega_{i,t}\right)\right)\left[1-\theta_{i,t}\left(\omega_{i,t}\right)\right]+l_i\left(e_{i,t}(\omega_{i,t})\right)\theta_{i,t}\left(\omega_{i,t}\right)\right\}dF_{i,t}\left(\omega_{i,t}\right).$$

We introduce the indicator functions $I_{i,t}^{\tilde{t}}(\omega_{i,t})$ which take value 1 if the length of the ongoing unemployment spell equals $\tilde{t}$ and 0 otherwise. Hence,

$$\begin{aligned}
Pr\left(I_{i,t+1}^{1}=1\right)&=\int l_i\left(e_{i,t}(\omega_{i,t})\right)\theta_{i,t}\left(\omega_{i,t}\right)dF_{i,t}\left(\omega_{i,t}\right),\\
Pr\left(I_{i,t+1}^{\tilde{t}}=1\right)&=\int\left(1-h_i\left(s_{i,t}\left(\omega_{i,t}\right),\omega_{i,t}\right)\right)I_{i,t}^{\tilde{t}-1}\left(\omega_{i,t}\right)dF_{i,t}\left(\omega_{i,t}\right).
\end{aligned}$$

The budget constraint still depends on the survival rate at each unemployment duration $S_{\tilde{t}}^{r}$, but now potentially spread over multiple spells. That is,

$$S_{\tilde{t}}^{r}=\Sigma_{t=1}^{T}\left[1+r\right]^{-(t-1)}\int\int I_{i,t}^{\tilde{t}}\left(\omega_{i,t}\right)dF_{i,t}\left(\omega_{i,t}\right)di.$$

Hence, the optimal formulae in Proposition 1 remain exactly the same (with the marginal utility of consumption when employed depending on effort on the job). The policy-relevant elasticity should account for potential responses in the layoff rate to a change in the unemployment policy. In our context, however, we find no significant responses in the layoff rates to changes in UI benefits.[2]

We refer to Chetty [2006] for a detailed treatment of other extensions of the model (including private insurance arrangements, spousal labour supply, etc.) which do not affect the optimal tax formulae due to envelope conditions. This remains true when extending his analysis to a dynamic benefit profile. For example, we can introduce alternative sources of income $z_{i,t}(x_{i,t},\omega_{i,t})$ into the agent's budget constraints (1) and/or (2), with the income level depending on the agent's choice variable $x_{i,t}$, which may enter the agent's utility function when employed and/or unemployed. As long as there are no externalities related to this alternative source of income, envelope conditions imply that the

[2] First, if layoff rates respond to the unemployment policy, this has implications regarding the pdf of daily wages around the kink in our empirical setting. The presence of a kink in benefits should create bunching at the kink if there is moral hazard on the job with convex costs of shirking. We show in subsection B.4 of Appendix B that we cannot detect any bunching at the kink. Furthermore, if layoffs are responsive to UI benefits this should also affect the pdf of daily wages when the kink in the schedule is removed. We show in subsection B.4 of Appendix B that we cannot detect such changes in the pdf of daily wages after the removal of the kinks in the schedule. While this evidence is far from definitive, it suggests that layoff rates do not seem to strongly respond to UI benefits in our context.

welfare impact of a policy change is still captured by the same statistics.

A.2.3 Characterization with Employer Screening

We consider a reduced-form model of employer screening based on Lockwood [1991] and Lehr [2017]. The job finding rate rate $h_{i,t}(s_{i,t}, \mathbf{S}_t) = \lambda_i(s_{i,t}) \times \mu_t(\mathbf{S}_t)$ is determined not only by the probability that agent i finds a vacant position $\lambda_i(s_{i,t})$, depending on her own search effort, but also by the probability that the matching firm hires the agent $\mu_t(\mathbf{S}_t)$ where $\mathbf{S}_t = \{S_{i',t}\}_{i'}$. The hiring probability at time t depends on the relative survival rates by agents with different productivity. In a two-type version of the model (H and L), where an agent i's type affects both her productivity θ_i and the probability of finding a vacancy $\lambda_i(s_{i,t})$, the firm's optimal hiring rate when matched with a job seeker who has been unemployed for t periods increases in the relative survival rate of the high type at time t. In particular, if the firm's profit of hiring an agent equals $\theta_i - w$, where $\theta^H > \theta^L = 0$, the firm's optimal hiring decision equals $\mu_t = 1$ if $\frac{S_t^H}{S_t^L}\theta^H \geq w$ and 0 otherwise. As a consequence, with the more productive type leaving unemployment at a faster rate, the firm would not hire job seekers who have been unemployed for longer than $\bar{t}$ where $\frac{S_{\bar{t}}^H}{S_{\bar{t}}^L}\theta^H = w$.

In the employer screening model, an agent's search effort will affect the job finding probability of any other agent, positively or negatively depending on her type, but no agent internalizes this effect. For simplicity we focus on job seekers' welfare and ignore the impact on firms' profits. Note that the setup can in principle also encompass richer models with rationing (e.g., Michaillat [2012]) and employer ranking (e.g., Blanchard and Diamond [1994]), in which job seekers' search effort crowd out the job finding rate of other job seekers. This is analyzed in Landais et al. [2010] who account for firms' profits and labor-demand behavior more generally and show how the distinction between micro and macro elasticities becomes relevant for the characterization of the optimal (static) unemployment policy in general equilibrium.

The impact of a policy change on the agents' welfare equals

$$\int \frac{\partial V_i(P)}{\partial b_t} di = \int S_{i,t} \frac{\partial v_i^u(c_{i,t}^u, s_{i,t})}{\partial c_{i,t}^u} di + \Sigma_{t'=1}^{\infty} [\int S_{i,t'} \lambda_i(s_{i,t'}) \frac{\partial \mu_{t'}}{\partial \mathbf{S}_{t'}} \frac{\partial \mathbf{S}_{t'}}{\partial b_t} \beta^{t'} \left[V_{i,t'}^e - V_{i,t'}^u\right] di]$$
$$= \lambda S_t \{[1 + CS_t] + \Sigma_{t'=1}^{\infty} \frac{S_{t'}}{S_t} [\int \frac{S_{i,t'}}{S_{t'}} \frac{\partial \lambda_i(s_{i,t'})\mu_{t'}}{\partial \mathbf{S}_{t'}} \frac{\partial \mathbf{S}_{t'}}{\partial b_t} \beta^{t'} \left[V_{i,t'}^e - V_{i,t'}^u\right] / \lambda di]\}$$
$$\equiv \lambda S_t \left\{[1 + CS_t] + \Sigma_{t'=1}^{\infty} \frac{S_{t'}}{S_t} E_{t'}^u \left(\frac{\partial h_{t'}}{\partial b_t} \omega_{t'}^h\right)\right\}.$$

We can correct the moral hazard cost for this new externality so that

$$\frac{\partial W(P)}{\partial b_t} = \lambda S_t \times [CS_t - MH_t^x],$$

with

$$MH_t^x \equiv \Sigma_{t'=1}^{T} \frac{S_{t'}}{S_t} \left[\frac{b_{t'} + \tau}{b_t} \varepsilon_{t',t} - E_{t'}^u \left(\omega_{t'}^h \frac{\partial h_{t'}}{\partial b_t}\right)\right],$$

The welfare impact of an increase in the exit rate rate is positive, $\omega_{i,t'}^h = \beta^{t'}\left[V_{i,t'}^e - V_{i,t'}^u\right]/\lambda$. The change in the exit rate $\partial h_{t'}/\partial b_t$ (through the change in hiring) will depend on the change in the relative survival rates at time t' in response to the change in benefits at time t. This corresponds to the correction proposed by Lehr [2017] for a flat benefit profile. A change in the unemployment policy won't change hiring, if the survival rate response of types with different productivity responds is the same. That is, $\varepsilon_{t',t}^i = \varepsilon_{t',t}^{i'}$.

Embedding this in our framework allows to assess the impact on the benefit profile as well. For the hiring externatity to change the gradient of the moral hazard costs, we need a change in the benefit level b_t to cause a different response in the relative survival rate (scaled by $1/S_t$) depending on the timing of the change. In our stylized two-type example the response in relative survival rate $S_{\bar{t}}^H/S_{\bar{t}}^L$ at the threshold duration $\bar{t}$ determines whether hiring increases or decreases. Indeed, the firm will hire job seekers with longer unemployment duration than $\bar{t}$ if the productive type is more responsive than the less productive type to a change in b_t, i.e., $\varepsilon_{\bar{t},t}^H > \varepsilon_{\bar{t},t}^L$. The externality response would be positive and thus causes $MH_t^x < MH_t$. In the model with heterogeneity in the returns to search, discussed in subsection 5.2.1 and considered below, we show that $\varepsilon_{\bar{t},t}^i$ can be increasing in the return to search for benefit levels paid early in the spell, but at the same decreasing for benefit levels paid late in the spell. Intuitively, the increase in the returns to search reduces the survival rate into longer unemployment spells and thus reduces the responsiveness to changes in benefits timed later on. Hence, with heterogeneity in the returns to search, the gradient of the moral hazard cost could become steeper when adjusted for the hiring externality.

A.2.4 Characterization with Income Taxation

We briefly illustrate the role of other fiscal externalities beyond the one introduced by the unemployment policy. In previous work on the Baily-Chetty formula, the only tax distortion in the economy comes from the unemployment policy. That is, no other revenue requirement exists ($\bar{G} = 0$) and the government imposes a lump sum contribution τ on the employed to balance the UI expenditures. Our model allows for taxes to fund an additional revenue requirement $\bar{G} > 0$. In practice, however, general government expenditures are funded through an income tax that is levied on both the employed and the unemployed.[3] Consider the case with a proportional income tax τ^y in addition to a lump sum UI contribution τ^u paid by employed workers. The (integrated) government budget can be rewritten as

$$G(P) - \bar{G} = [T - D](\tau^u + \tau^y w) - \Sigma S_t (b_t - \tau^y b_t) - \bar{G},$$

where

$$\begin{aligned}
\frac{\partial G(P)}{\partial b_t} &= -S_t^r(1-\tau^y) - \Sigma_{t'=1}^T (b_{t'} - \tau^y b_{t'} + \tau^u + \tau^y w)\frac{\partial S_{t'}^r}{\partial b_t} \\
&= -S_t^r(1-\tau^y) - \Sigma_{t'=1}^T (b_{t'} + \tau^u)\frac{\partial S_{t'}^r}{\partial b_t} - \Sigma_{t'=1}^T \tau^y (w - b_{t'})\frac{\partial S_{t'}^r}{\partial b_t} \\
&= -D_k^r\left[(1-\tau^y) + \Sigma_{t'=1}^T \frac{S_{t'}^r}{S_t^r}\frac{b_{t'}+\tau^u}{b_t}\varepsilon_{t',t}^r + \Sigma_{t'=1}^T \frac{S_{t'}^r}{S_t^r}\frac{\tau^y (w - b_{t'})}{b_t}\varepsilon_{t',t}^r\right].
\end{aligned}$$

The first two terms capture the standard mechanical and behavioral effect of an increase in the benefit level on the expenditures and revenues related to the unemployment policy. The third term captures the fiscal externality through the income tax, accounting for the reduction in income tax revenues when increasing unemployment. For a flat profile, this effect is proportional to $\tau^y \frac{w-\bar{b}}{\bar{b}}$ and thus small when the average effective income tax rate is small or the replacement rate is high. It is a standard simplification in related work to ignore these fiscal spillover effects across different government policies. Note also that from the consumption smoothing perspective, the difference in marginal utilities remains sufficient.

A.2.5 Welfare Impact of Change in Tilt

Corollary 1. *Whenever $\frac{CS_1}{MH_1} > \frac{CS_2}{MH_2}$, welfare can be increased by increasing the tilt b_1/b_2. A budget-balanced increase in the tilt b_1/b_2 increases welfare if and only if $\frac{1+CS_1}{1+MH_1} > \frac{1+CS_2}{1+MH_2}$.*

[3]In Sweden, UI benefits are fully included in individuals' taxable income to the personal income tax.

Proof: *By implicit differentiation, we find that when increasing $b_{\mathbf{1}}$ and decreasing $b_{\mathbf{2}}$ at rate*

$$\frac{db_{\mathbf{2}}}{db_{\mathbf{1}}}|_1 = -\frac{D_{\mathbf{1}}\left(1+MH_{\mathbf{1}}\right)}{D_{\mathbf{2}}\left(1+MH_{\mathbf{2}}\right)}, \tag{3}$$

the policy budget remains balanced. The welfare impact of this budget-balanced increase in the tilt $b_{\mathbf{1}}/b_{\mathbf{2}}$ equals

$$\begin{aligned}\frac{\partial W\left(P\right)}{\partial b_{\mathbf{1}}} - \frac{\partial W\left(P\right)}{\partial b_{\mathbf{2}}}\frac{db_{\mathbf{2}}}{db_{\mathbf{1}}}|_1 &= \lambda D_{\mathbf{1}}\left[1+CS_{\mathbf{1}}-1-MH_{\mathbf{1}}\right] - \lambda D_{\mathbf{2}}\left[1+CS_{\mathbf{2}}-1-MH_{\mathbf{2}}\right]\frac{D_{\mathbf{1}}\left(1+MH_{\mathbf{1}}\right)}{D_{\mathbf{2}}\left(1+MH_{\mathbf{2}}\right)} \\ &= \lambda D_{\mathbf{1}}\left(1+MH_{\mathbf{1}}\right)\times\left\{\frac{1+CS_{\mathbf{1}}}{1+MH_{\mathbf{1}}} - \frac{1+CS_{\mathbf{2}}}{1+MH_{\mathbf{2}}}\right\}.\end{aligned}$$

This proves the second part of the corollary. Consider now an increase in $b_{\mathbf{1}}$ jointly with a decrease in $b_{\mathbf{2}}$ at rate

$$\frac{db_{\mathbf{2}}}{db_{\mathbf{1}}}|_2 = -\frac{D_{\mathbf{1}}MH_{\mathbf{1}}}{D_{\mathbf{2}}MH_{\mathbf{2}}}$$

The welfare impact of such increase in the tilt $b_{\mathbf{1}}/b_{\mathbf{2}}$ equals

$$\begin{aligned}\frac{\partial W\left(P\right)}{\partial b_{\mathbf{1}}} - \frac{\partial W\left(P\right)}{\partial b_{\mathbf{2}}}\frac{db_{\mathbf{2}}}{db_{\mathbf{1}}}|_2 &= \lambda D_{\mathbf{1}}\left[CS_{\mathbf{1}}-MH_{\mathbf{1}}\right] - \lambda D_{\mathbf{2}}\left[CS_{\mathbf{2}}-MH_{\mathbf{2}}\right]\frac{D_{\mathbf{1}}MH_{\mathbf{1}}}{D_{\mathbf{2}}MH_{\mathbf{2}}} \\ &= \lambda D_{\mathbf{1}}MH_{\mathbf{1}}\times\left\{\frac{CS_{\mathbf{1}}}{MH_{\mathbf{1}}} - \frac{CS_{\mathbf{2}}}{MH_{\mathbf{2}}}\right\}.\end{aligned}$$

Hence, whenever $CS_{\mathbf{1}}/MH_{\mathbf{1}}$ exceeds $CS_{\mathbf{2}}/MH_{\mathbf{2}}$, such increase in the tilt $b_{\mathbf{1}}/b_{\mathbf{2}}$ increases welfare and vice versa.■

A.3 Dynamic Sufficient Statistics in Stationary Environment

A.3.1 Proof of Proposition 2

We consider a flat benefit profile $b_t = \bar{b} < w - \tau$ for $\forall t$ in a single-type, stationary environment $h_{i,t}\left(\cdot\right) = \bar{h}\left(\cdot\right)$ for $\forall i, t$. We also assume $\beta\left(1+r\right) = 1$ and $T = \infty$. We compare the impact of an increase in the benefit level at time t and at time $t+1$.

We analyze first the moral hazard costs. We assume that the agent is borrowing constrained and thus consumes hand-to-mouth when unemployed and employed ($c_t^u = b_t$ and $c_t^e = w - \tau$). This set up follows Hopenhayn and Nicolini [1997]. Using notation $S_t^r = \left(1+r\right)^{-(t-1)}S_t$, we find

$$\begin{aligned}\frac{\partial G\left(P\right)}{\partial b_t} &= -S_t^r - \Sigma_{j=1}^T\left(b_j+\tau\right)\frac{\partial S_j^r}{\partial b_k} \\ &= -S_t^r\times\left[1+\frac{b+\tau}{b}\frac{D^r}{S_t^r}\varepsilon_{D^r,b_t}\right].\end{aligned}$$

For an increase in b_{t+1}, we find

$$\frac{\partial G\left(P\right)}{\partial b_{t+1}} = -S_{t+1}^r\times\left[1+\frac{b+\tau}{b}\frac{D^r}{S_{t+1}^r}\varepsilon_{D^r,b_{t+1}}\right].$$

Using

$$D^r = \Sigma_{j=1}^T S_j^r = 1 + D_{\mathbf{2}}^r = 1 + S_2^r\tilde{D}_{\mathbf{2}}^r,$$

where $D_{\mathbf{2}}^r = \Sigma_{j=2}^T S_j^r$ and $\tilde{D}_{\mathbf{2}}^r = \left[\Sigma_{j=2}^T S_j^r/S_2^r\right]$, we can write

$$\begin{aligned}\varepsilon_{D^r,b_{t+1}} &= \frac{\partial\left[1+D_{\mathbf{2}}^r\right]}{\partial b_{t+1}}\frac{b}{D^r} = \frac{\partial D_{\mathbf{2}}^r}{\partial b_{t+1}}\frac{b}{D_{\mathbf{2}}^r}\frac{D_{\mathbf{2}}^r}{D^r} \\ &= \left[\varepsilon_{S_2^r,b_{t+1}} + \varepsilon_{\tilde{D}_{\mathbf{2}}^r,b_{t+1}}\right]\frac{D_{\mathbf{2}}^r}{D^r}.\end{aligned}$$

Since the environment is stationary and the agent is borrowing-constrained, the agent's search behavior remains the *same over the unemployment spell (conditional on the continuation policy being the same). Starting from a flat profile, an increase in b_t has the same impact on the continuation policy evaluated at time 1 as an increase in b_{t+1} has on the continuation policy evaluated at time 2, conditional on being still unemployed then. The impact of the policy changes at time t and $t+1$ on the remaining duration at time 1 and time 2 respectively is the same. Hence, we have $\varepsilon_{\tilde{D}_{\mathbf{2}}^r,b_{t+1}} = \varepsilon_{D^r,b_t}$ for $T = \infty$. Denoting the constant exit rate for the flat profile by h, we have $D^r = \frac{1+r}{r+h}$ and $D_{\mathbf{2}}^r = \frac{1-h}{1+r}\frac{1+r}{r+h}$, while $S_{t+1}^r = \frac{1-h}{1+r}S_t^r$. This implies*

$$\frac{D_{\mathbf{2}}^r}{S_{t+1}^r} = \frac{D^r}{S_t^r}.$$

Using this equality and the expression for $\varepsilon_{D^r,b_{t+1}}$, we can re-write

$$\begin{aligned}\frac{\partial G(P)}{\partial b_{t+1}} &= -S_{t+1}^r \times [1 + \frac{b+\tau}{b}\frac{D^r}{S_{t+1}^r}\varepsilon_{D^r,b_{t+1}}] \\ &= -S_{t+1}^r \times [1 + \frac{b+\tau}{b}\frac{D_{\mathbf{2}}^r}{S_{t+1}^r}\left[\varepsilon_{S_2^r,b_{t+1}} + \varepsilon_{D^r,b_t}\right]] \\ &= -S_{t+1}^r \times [1 + \frac{b+\tau}{b}\frac{D^r}{S_t^r}\left[\varepsilon_{S_2^r,b_{t+1}} + \varepsilon_{D^r,b_t}\right]].\end{aligned}$$

This implies that

$$MH_{t+1} = \frac{b+\tau}{b}\frac{D^r}{S_t^r}\varepsilon_{S_2^r,b_{t+1}} + \frac{b+\tau}{b}\frac{D^r}{S_t^r}\varepsilon_{D^r,b_t} \geq MH_t,$$

since $\varepsilon_{S_2^r,b_{t+1}} \geq 0$. Starting from a flat profile, the moral hazard cost is thus higher for any benefit increase that is timed later during the spell.

We now analyze the consumption smoothing gains. In our stylized setup (which assumes that the agent starts unemployed and remains employed once she finds a job), an optimizing agent's unemployment consumption $c_t^u(\omega_t)$ (and search effort $s_t(\omega_t)$) only depends on the length of the ongoing unemployment spell. Hence, we have

$$\int \frac{\partial V_i(P)}{\partial b_t}di = \beta^{t-1}S_t\frac{\partial v^u(c_t^u, s_t)}{\partial c_t^u}.$$

When the agent is borrowing constrained, the agent is hand-to-mouth $c_t^u = b_t$ and the marginal utility of consumption (and thus CS_t) remains constant for a flat benefit profile. When not borrowing constrained, an agent who is unemployed at time t increases her consumption by depleting her assets to equalize the marginal utility of consumption at time t with the expected marginal utility of consumption at $t+1$. The unemployment consumption level c_t^u at time t, the consumption level upon finding employment c_{t+1}^e at time $t+1$ and the consumption level when still being unemployed c_{t+1}^u at time $t+1$ satisfy a standard Euler condition,

$$\frac{\partial v^u(c_t^u, s_t)}{\partial c_t^u} = h_t(s_t)\frac{\partial v^e(c_{t+1}^e)}{\partial c_{t+1}^e} + (1 - h_t(s_t))\frac{\partial v^u(c_{t+1}^u, s_{t+1})}{\partial c_{t+1}^u}$$

for $\beta(1+r) = 1$. With separable concave preferences, $\partial v^u(c,s)/\partial c = \partial v^e(c)/\partial c = v'(c)$, and, benefits lower than the after-tax wage $b < w - \tau$, for any given asset level, an agent has higher expected lifetime income when employed than when unemployed. The marginal value of an increase in assets is lower when employed than when unemployed, i.e., $\partial V_{t+1}^e/\partial a_{t+1} < \partial V_{t+1}^u/\partial a_{t+1}$. This implies the marginal utility of consumption is lower when employed than when unemployed at $t+1$. Hence, by the Euler condition, $v'(c_{t+1}^u) > v'(c_{t+1}^e)$ implies $v'(c_{t+1}^u) > v'(c_t^u)$. On the optimal path, the marginal utility of consumption is increasing over the spell and the consumption gains are thus always higher for benefits timed later during the unemployment spell.■

The stationary forces and how they affect the optimal benefit profile are well known in the literature and arguably robust. Our set up with the borrowing-constrained agent follows Hopenhayn and Nicolini [1997]. The assumption that the agent is borrowing constrained is restrictive, but guarantees that search behaviour remains the same over the unemployment spell and thus simplifies the derivations. Note that search behaviour remains the same in a model with savings when the agent has CARA preferences with monetary cost of search efforts (i.e.,

$v^u(c,s) = -\exp(-\sigma[c-\psi(s)])$ as in Spinnewijn [2015], again simplifying the derivation of the optimal benefit profile. It is also clear from the proof that relaxing the borrowing constraint would not change the conclusion regarding the gradient of the moral hazard costs when $\varepsilon_{\tilde{D}^r_{\mathbf{2}},b_{t+1}} \geq \varepsilon_{D^r,b_t}$ and $h_{t+1} \geq h_t$ (so that $\frac{D^r_{\mathbf{2}}}{S^r_{t+1}} > \frac{D^r}{S^r_t}$) for any t. We analyze this further for a specific search environment with non-stationary features. The result that the marginal utility of consumption is increasing over the spell continues to hold for unconstrained job seekers when the benefit profile is not flat but $b_t < w - \tau$ for all t. The assumption that the agent's preferences are separable is also more restrictive than necessary. The proof highlights that it suffices for the marginal value of an increase in assets to be lower when employed than when unemployed.

A.4 Dynamic Sufficient Statistics in a Non-stationary Environment

We now specify particular functions for the search environment and introduce non-stationary features in our model. We allow for depreciation in search efficacy, heterogeneity in search efficacy, and heterogeneity in assets. We study how these forces affect the predicted increase in MH_t and CS_t throughout the unemployment spell from Proposition 2. We allow for all these non-stationary forces simultaneously in our structural model in Appendix D.

We establish three results: (i) in a model with depreciation in the return-to-search parameter (i.e., $h_t(s_{i,t}) = h_0 + \theta^t h_1 s^\rho_{i,t}$), the moral hazard cost cost of benefit changes that start later in the spell can be arbitrarily close to the moral hazard cost of benefit changes that start earlier, (ii) in a model with heterogeneity in the return-to-search parameter (i.e., $h_t(s_{i,t}) = h_0 + h^i_1 s^\rho_{i,t}$), the moral hazard costs can actually be lower for benefit changes timed later in the spell,, (iii) in a model with asset heterogeneity, negatively correlated with exit rates, the consumption smoothing gains can actually be higher earlier in the spell.

A.4.1 Preliminaries

For our analysis of moral hazard costs, we assume that agents are borrowing constrained throughout the unemployment spell (i.e., unemployment consumption equals UI benefits), that preferences are separable in consumption and search, $u(c,s) = u(c) - s$, and that the exit rate function has the following form,

$$h(s_{i,t}) = h_0 + h_1 s^\rho_{i,t} \text{ for } \forall i,t.$$

For tractability, we assume that the optimal search effort is interior and thus the resulting exit rate is between 0 and 1.

Each individual has a value function for the employed and unemployed state shown below:

$$\begin{aligned} V^e_{i,t} &= u(w-\tau) + \beta V^e_{i,t+1} \\ V^u_{i,t} &= u(b_t) - s_{i,t} + \beta h_{i,t}(s_{i,t})[V^e_{i,t+1} - V^u_{i,t+1}] + \beta V^u_{i,t+1}, \end{aligned}$$

Since the employment state is absorbing, we have $V^e_t = \frac{u(w-\tau)}{1-\beta}$.The optimal level of effort equals

$$s_{i,t} = (\rho\beta h_{1,i,t}[V^e_{i,t+1} - V^u_{i,t+1}])^{\frac{1}{1-\rho}}.$$

We start from a flat benefit profile and compare a permanent benefit rise in $t=2$ (denoted by $b_{2\to\infty}$) and in $t=3$ (denoted by $b_{3\to\infty}$) respectively. Note that $t=1$ is the first period that an agent exerts effort, but this is unaffected by the benefit level b_1. The moral hazard cost of raising benefits permanently in period t, starting from a flat profile, is given by:

$$MH_{t\to\infty} = \frac{\frac{\partial D^r}{\partial b_{t\to\infty}}}{D^r_{t\to\infty}}(b+\tau).$$

To save on notation we consider instead $D = \sum_{t'=1}^\infty S_{t'}$ and $D_{t\to\infty} = \sum_{t'=t}^\infty S_{t'}$, corresponding to D^r and $D^r_{t\to\infty}$ for $r=0$, but we make sure our conclusions are robust to $\beta \to 1$.

A.4.2 Stationary Environment

We first confirm that for this specific search environment, in the absence of non-stationary features, the moral hazard cost of increasing the UI benefits is always higher when this increase is timed later in the spell, in line with Proposition 2. This will help highlighting why non-stationary features can affect this result.

Consider first an increase in $b_{2\to\infty}$. In this scenario, we have that $V_t^u = V_{t+1}^u = V_2^u$ if $t \geq 2$. As a result we have only two value functions when unemployed:

$$\begin{aligned} V_1^U &= u(b) - s + \beta(h_0 + h_1 s^\rho)[V^e - V_2^u] + \beta V_2^u \\ V_2^u &= u(b_{2\to\infty}) - s + \beta(h_0 + h_1 s^\rho)[V^e - V_2^u] + \beta V_2^u \end{aligned}$$

and one level of effort:

$$s = (\rho\beta h_1[V^e - V_2^u])^{\frac{1}{1-\rho}}.$$

Hence, we can write

$$\begin{aligned} S_t &= (1 - h_0 - h_1 s^\rho)^{t-1}, \\ D &= \frac{1}{h_0 + h_1 s^\rho} \\ D_{2\to\infty} &= \frac{1 - h_0 - h_1 s^\rho}{h_0 + h_1 s^\rho} \end{aligned}$$

Since we evaluate the benefit change for a flat profile, we will use the fact that before differentation $V_1^u = V_2^u = V^u$. We calculate the effect of the benefit rise on the average unemployment duration, which in turn depends on the change in effort, which in turn depends on the change in the value of being unemployed:

$$\begin{aligned} \frac{\partial D}{\partial b_{2\to\infty}} &= -\frac{\rho h_1 s^{\rho-1} \frac{\partial s}{\partial b_{2\to\infty}}}{(h_0 + h_1 s^\rho)^2}, \\ \frac{\partial s}{\partial b_{2\to\infty}} &= -\frac{s}{1-\rho}[V^e - V^u]^{-1} \frac{\partial V_2^u}{\partial b_{2\to\infty}}, \\ \frac{\partial V_2^u}{\partial b_{2\to\infty}} &= \frac{u'(\cdot)}{1 - \beta(1 - h_0 - h_1 s^\rho)}. \end{aligned}$$

Consider now an increase in $b_{3\to\infty}$. Note that $V_t^u = V_{t+1}^u = V_3^U$ if $t \geq 3$. Therefore, there are only three value functions when unemployed:

$$\begin{aligned} V_1^u &= u(b) - s_1 + \beta(h_0 + h_1 s_1^\rho)[V^e - V_2^u] + \beta V_2^u \\ V_2^u &= u(b) - s_2 + \beta(h_0 + h_1 s_2^\rho)[V^e - V_3^u] + \beta V_3^u \\ (1-\beta)V_3^U &= u(b_{3\to\infty}) - s_2 + \beta(h_0 + h_1 s_2^\rho)[V^e - V_3^u], \end{aligned}$$

and two levels of effort:

$$\begin{aligned} s_1 &= (\rho\beta h_1[V^e - V_2^u])^{\frac{1}{1-\rho}} \\ s_2 &= (\rho\beta h_1[V^e - V_3^u])^{\frac{1}{1-\rho}}. \end{aligned}$$

Similar to before, we find

$$\frac{\partial D}{\partial b_{3\to\infty}} = \frac{-\rho h_1 s^{\rho-1} \frac{\partial s_1}{\partial b_{3\to\infty}}(h_0 + h_1 s^\rho) - \rho h_1 s^{\rho-1} \frac{\partial s_2}{\partial b_{3\to\infty}}(1 - h_0 - h_1 s^\rho)}{(h_0 + h_1 s^\rho)^2},$$

which is composed of the following derivatives:

$$\frac{\partial s_1}{\partial b_{3\to\infty}} = -\frac{s}{1-\rho}[V^e - V^u]^{-1}\frac{\partial V_2^u}{\partial b_{3\to\infty}}$$
$$\frac{\partial s_2}{\partial b_{3\to\infty}} = -\frac{s}{1-\rho}[V^e - V^u]^{-1}\frac{\partial V_3^u}{\partial b_{3\to\infty}},$$

which are, in turn, composed of the following derivatives:

$$\frac{\partial V_3^U}{\partial b_{3\to\infty}} = \frac{u'(\cdot)}{1-\beta(1-h_0-h_1 s^\rho)},$$
$$\frac{\partial V_2^U}{\partial b_{3\to\infty}} = -\beta(h_0+h_1 s^\rho)\frac{\partial V_3^u}{\partial b_{3\to\infty}} + \beta\frac{\partial V_3^u}{\partial b_{3\to\infty}}.$$

Putting everything together, we find for $b_{2\to\infty}$ and $b_{3\to\infty}$:

$$\frac{\partial D}{\partial b_{3\to\infty}} = \frac{\partial D}{\partial b_{2\to\infty}}(1-h_0-h_1 s^\rho)[1+\beta(h_0+h_1 s^\rho)],$$

while

$$D_{3\to\infty} = D_{2\to\infty}(1-h_0-h_1 s^\rho).$$

The impact of an increase in $b_{3\to\infty}$ on the time spent unemployed is smaller than the impact of an increase in $b_{2\to\infty}$, since a smaller part of the unemployment policy is affected, as captured by the scalar $[1-h_0-h_1 s^\rho]$ in the both expressions above, for the duration responses and the durations respectively. However, while the increase in $b_{3\to\infty}$ starts later, it will reduce the exit rates earlier in the spell as well, as captured by the scalar $[1+\beta(h_0+h_1 s^\rho)]$ in the expression for the duration responses. These are the forward-looking incentives identified before in Shavell and Weiss [1979]. Indeed, in line with Proposition 2, we find

$$\frac{MH_{2\to\infty}}{MH_{3\to\infty}} = \frac{1}{1+\beta(h_0+h_1 s^\rho)} < 1.$$

This intuition generalizes for any changes $b_{t\to\infty}$ and $b_{t+1\to\infty}$ respectively and is robust to $\beta \to 1$.

A.4.3 Depreciation in Search Efficacy

We now assume that the returns to search depreciate at a geometric rate,

$$h_t\left(s_{i,t}\right) = h_0 + \theta^{t-1} h_1 s_{i,t}^\rho = \theta^{t-1} h s_{i,t}^\rho \text{ for } \theta \in [0,1].$$

To simplify the expressions below, we assume that the exit rate is zero when no search is exerted (i.e., $h_0 = 0$), but the argument below continues to apply when relaxing this assumption. From the value of being unemployed at time t and the effort level at time t, we can derive the following derivatives:

$$\frac{\partial V_t^u}{\partial b_{\bar{t}\to\infty}} = \beta(1-h\theta^{t-1}s_t^\rho)\frac{\partial V_{t+1}^u}{\partial b_{\bar{t}\to\infty}} \;\forall\; 0 < t < \bar{t}$$
$$\frac{\partial V_t^u}{\partial b_{\bar{t}\to\infty}} = u'\left(b\right) + \beta(1-h\theta^{t-1}s_t^\rho)\frac{\partial V_{t+1}^u}{\partial b_{\bar{t}\to\infty}} \;\forall\; t \geq \bar{t}$$
$$\frac{\partial s_t}{\partial b_{\bar{t}\to\infty}} = \frac{-s_t}{1-\rho}[V^e - V_{t+1}^u]^{-1}\frac{\partial V_{t+1}^u}{\partial b_{\bar{t}\to\infty}},$$

which can be used in order to derive an expression for the derivative of the average unemployment duration:

$$
\begin{aligned}
\frac{\partial D}{\partial b_{\bar{t}\to\infty}} &= \frac{\partial}{\partial b_{\bar{t}\to\infty}}[1+(1-hs_1^\rho)+(1-hs_1^\rho)(1-h\theta s_2^\rho)+(1-hs_1^\rho)(1-h\theta s_2^\rho)(1-h\theta^2 s_3^\rho)+\ldots] \\
&= -\rho h\Big[s_1^{\rho-1}\frac{\partial s_1}{\partial b_{\bar{t}\to\infty}}\frac{D_{2\to\infty}}{1-hs_1^\rho}+\theta s_2^{\rho-1}\frac{\partial s_2}{\partial b_{\bar{t}\to\infty}}\frac{D_{3\to\infty}}{1-h\theta s_2^\rho}+\ldots\Big] \\
&= -\rho h\sum_{t'=1}^{\infty}\theta^{t'-1}s_{t'}^{\rho-1}\frac{D_{t'+1\to\infty}}{1-h\theta^{t'-1}s_{t'}^\rho}\frac{\partial s_{t'}}{\partial b_{\bar{t}\to\infty}} \\
&= \frac{\rho}{1-\rho}h\sum_{t'=1}^{\infty}\theta^{t'-1}s_{t'}^{\rho}[V^e-V_{i+1}^u]^{-1}\frac{D_{t'+1\to\infty}}{1-h\theta^{t'-1}s_{t'}^\rho}\frac{\partial V_{t'+1}^u}{\partial b_{\bar{t}'\to\infty}}
\end{aligned}
$$

We then use the following feature:

$$\frac{\partial V_t^u}{\partial b_{\bar{t}\to\infty}} = \frac{\partial V_t^u}{\partial b_{\hat{t}\to\infty}} \;\forall\; t\geq \max[\bar{t},\hat{t}]$$

and

$$
\begin{aligned}
\frac{\partial V_2^u}{\partial b_{2\to\infty}} &= u'(\cdot)+\beta(1-h\theta s_2^\rho)\frac{\partial V_3^U}{\partial b_{2\to\infty}} \\
&= u'(\cdot)+\frac{\partial V_2^U}{\partial b_{3\to\infty}}
\end{aligned}
$$

to re-express the ratio of duration responses as

$$
\begin{aligned}
\frac{\frac{\partial D}{\partial b_{2\to\infty}}}{\frac{\partial D}{\partial b_{3\to\infty}}} &= \frac{s_1^\rho[V^e-V_2^u]^{-1}\frac{D_{2\to\infty}}{1-hs_1^\rho}\frac{\partial V_2^u}{\partial b_{2\to\infty}}+\sum_{t'=2}^{\infty}\theta^{t'-1}s_{t'}^\rho[V^e-V_{t'+1}^u]^{-1}\frac{D_{t'+1\to\infty}}{1-h\theta^{t'-1}s_{t'}^\rho}\frac{\partial V_{t'+1}^u}{\partial b_{2\to\infty}}}{s_1^\rho[V^e-V_2^u]^{-1}\frac{D_{2\to\infty}}{1-hs_1^\rho}\frac{\partial V_2^u}{\partial b_{3\to\infty}}+\sum_{t'=2}^{\infty}\theta^{t'-1}s_{t'}^\rho[V^e-V_{t'+1}^u]^{-1}\frac{D_{t'+1\to\infty}}{1-h\theta^{t'-1}s_{t'}^\rho}\frac{\partial V_{t'+1}^u}{\partial b_{3\to\infty}}} \\
&= \frac{u'(\cdot)+\frac{\partial V_2^u}{\partial b_{3\to\infty}}+A}{\frac{\partial V_2^u}{\partial b_{3\to\infty}}+A},
\end{aligned}
$$

where

$$A=\sum_{t'=2}^{\infty}\theta^{t'-1}\frac{s_{t'}^\rho}{s_1^\rho}\frac{V^e-V_2^u}{V^e-V_{t'+1}^U}\frac{D_{t'+1\to\infty}}{D_{2\to\infty}}\frac{1-hs_1^\rho}{1-h\theta^{t'-1}s_{t'}^\rho}\frac{\partial V_{t'+1}^u}{\partial b_{j\to\infty}} \text{ for } j=2,3,$$

capturing the response in exit rates from time 2 onwards. Hence, we can find an explicit expression for the ratio of moral hazard costs,

$$\frac{MH_{2\to\infty}}{MH_{3\to\infty}} = \frac{u'(\cdot)+\frac{\partial V_2^u}{\partial b_{3\to\infty}}+A}{\frac{\partial V_2^u}{\partial b_{3\to\infty}}+A}\frac{D_{3\to\infty}}{D_{2\to\infty}}.$$

Note that when we set $\theta=1$, we return to the stationary model and find that $MH_{2\to\infty}/MH_{3\to\infty}<1$. However, in this non-stationary setting, the ratio depends crucially on θ and can be made arbitrarily close to 1 for sufficiently low θ. That is, using $D_{3\to\infty}=D_{2\to\infty}-(1-hs_1^\rho)$,

$$
\begin{aligned}
MH_{2\to\infty}/MH_{3\to\infty} &\approx 1 \Leftrightarrow \\
u'(\cdot)[D_{2\to\infty}-(1-hs_1^\rho)] &\approx (1-hs_1^\rho)\Big[\frac{\partial V_2^u}{\partial b_{3\to\infty}}+A\Big] \Leftrightarrow \\
u'(\cdot)D_{3\to\infty} &\approx (1-hs_1^\rho)\Big[\beta(1-h\theta s_2^\rho)\frac{\partial V_3^u}{\partial b_{3\to\infty}}+A\Big] \Leftrightarrow \\
u'(\cdot)D_{3\to\infty} &\approx (1-hs_1^\rho)\Big[\beta(1-h\theta s_2^\rho)\Big(u'(\cdot)+\beta(1-h\theta^2 s_3^\rho)\frac{\partial V_4^U}{\partial b_{3\to\infty}}\Big)+A\Big]
\end{aligned}
$$

Iterating the substitution of $\frac{\partial V_t^u}{\partial b_{3\to\infty}}=u'(\cdot)+\beta(1-h\theta s_t^\rho)\frac{\partial V_{t+1}^U}{\partial b_{3\to\infty}}$, we find

$$u'(\cdot)D_{3\to\infty}\approx\Big[\beta(1-hs_1^\rho)(1-h\theta s_2^\rho)u'(\cdot)+\beta^2(1-hs_1^\rho)(1-h\theta s_2^\rho)(1-h\theta^2 s_3^\rho)u'(\cdot)+\ldots\Big]+(1-hs_1^\rho)A.$$

The bracketed term in the RHS converges to $u'(\cdot)D_{3\to\infty}$ for $\beta \to 1$. (In particular, when properly discounting the survival rates to calculate the moral hazard costs, the two terms would coincide for $\beta(1+r)=1$.) This shows the importance of the term A, determined by the exit rate responses later in the spell, in driving the wedge between the moral hazard costs. This wedge is still positive, like in the stationary model, since $A>0$. However, the wedge disappears when A converges to 0. Now by setting θ arbitrarily small we can make A arbitrarily small, since the terms in the summation are scaled by $\theta^{t'-1}$ and thus converge to 0, while all other factors can be bounded from above. Hence, for small enough θ, we have that $MH_{2\to\infty} \approx MH_{3\to\infty}$.

The intuition underlying this result is that depreciation in the returns to search reduces the responsiveness in the exit rates later in the unemployment spell to changes in the UI benefits. While this force cannot reverse the relative magnitude of the moral hazard cost, it mitigates the weight on the forward-looking incentives in driving this wedge. We now turn to a case where the relative magnitudes can actually be reversed.

A.4.4 Heterogeneity in Search Efficacy

We now consider heterogeneity in search efficacy, allowing for two types of agents, type y and type z. Type-y agents have higher return to their search effort,

$$h_1^y > h_1^z.$$

The proportion of y-types at the start of the unemployment spell equals α.

Our approach is different from the stationary case and the case with search depreciation in which we derived an explicit expression for $(\partial D/\partial b_{t\to\infty})/D_{t\to\infty}$. Instead we follow the approach in the proof of Proposition 2 and Section 5.2. We decompose the moral hazard cost of raising benefits in period 3 permanently into the response to forward-looking incentives and the response in the *remaining duration* of unemployment, conditional on still being unemployed in period 3,

$$MH_{3\to\infty} \times \frac{b}{b+\tau} = \frac{D_{1\to 2}}{D_{3\to\infty}} \varepsilon_{D_{1\to 2},b_{3\to\infty}} + \varepsilon_{S_3,b_{3\to\infty}} + \varepsilon_{\tilde{D}_{3\to\infty},b_{3\to\infty}}, \tag{4}$$

where $\tilde{D}_{3\to\infty} \equiv \frac{D_{3\to\infty}}{S_3}$ and $D_{1\to 2} \equiv S_1 + S_2$. In a single-agent model without heterogeneity, the latter response corresponds to the moral hazard cost of an overall increase in benefits, $\varepsilon_{\tilde{D}_{3\to\infty},b_{3\to\infty}} = \varepsilon_{D,b_{1\to\infty}}$. With heterogeneity, the magnitude and the weights attached to the different elasticities depend on the different y- and z-types and their respective survival.

We first show that, for a given type, all three terms in (4) are increasing in search efficacy h_1. Define the following common component amongst all three terms

$$B = \frac{\rho}{1-\rho} \frac{u'(b)}{(1-\beta(1-h_0-h_1 s^\rho))} \frac{b}{[V^e - V^u]},$$

We then have

$$\begin{aligned}
\varepsilon_{\tilde{D}_{3\to\infty},b_{3\to\infty}} = \varepsilon_{D,b_{1\to\infty}} &= B\frac{h_1 s^\rho}{h_0 + h_1 s^\rho} \\
\varepsilon_{S_3,b_{3\to\infty}} &= B\frac{h_1 s^\rho[1+\beta(1-h_0-h_1 s^\rho)]}{1-h_0-h_1 s^\rho} \\
\frac{D_{1\to 2}}{D_{3\to\infty}}\varepsilon_{D_{1\to 2},b_{3\to\infty}} &= B\frac{h_1 s^\rho \beta(h_0+h_1 s^\rho)}{1-h_0-h_1 s^\rho}.
\end{aligned}$$

For tractability, we continue under the assumption that

$$\frac{\partial B}{\partial h_1} = -B\cdot\frac{\rho}{1-\rho}\frac{\beta s^\rho(1-\beta)}{(1-\beta(1-h_0-h_1 s^\rho))^2} \approx 0,$$

which follows from $1-\beta \approx 0$. We also have

$$\frac{\partial\left[h_0+h_1 s^\rho\right]}{\partial h_1} = s^\rho + h_1 \rho s^{\rho-1}\frac{\partial s}{\partial h_1} = s^\rho + \frac{\rho(1-\beta)s^\rho}{(1-\rho)(1-\beta(1-h_0-h_1 s^\rho))} \approx s^\rho,$$

using again $1-\beta \approx 0$. As a consequence, all the above terms are increasing in search efficacy h_1.

The above elasticities are derived for a given type. This aggregrates up as follows:

$$\begin{aligned}
\varepsilon_{\bar{D},b_{1\to\infty}} &= \alpha\frac{D^y}{\bar{D}}\varepsilon_{D^y,b_{1\to\infty}} + (1-\alpha)\frac{D^z}{\bar{D}}\varepsilon_{D^z,b_{1\to\infty}},\\
\varepsilon_{\bar{S}_3,b_{3\to\infty}} &= \alpha\frac{D^y_{3\to\infty}}{\bar{D}_{3\to\infty}}\varepsilon_{S_3^y,b_{3\to\infty}} + (1-\alpha)\frac{D^z_{3\to\infty}}{\bar{D}_{3\to\infty}}\varepsilon_{S_3^z,b_{3\to\infty}},\\
\frac{\bar{D}_{1\to 2}}{\bar{D}_{3\to\infty}}\varepsilon_{\bar{D}_{1\to 2},b_{3\to\infty}} &= \alpha\frac{D^y_{1\to 2}}{\bar{D}_{3\to\infty}}\varepsilon_{D^y_{1\to\infty},b_{3\to\infty}} + (1-\alpha)\frac{D^z_{1\to 2}}{\bar{D}_{3\to\infty}}\varepsilon_{D^z_{1\to\infty},b_{3\to\infty}},
\end{aligned}$$

To emphasize the difference, we have introduced the upper-bar notation to refer to aggregates. We now wish to show that $MH_{1\to\infty} > MH_{3\to\infty}$ is true in the presence of sufficient heterogeneity. This is equivalent to:

$$\varepsilon_{\bar{D},b_{1\to\infty}} > \frac{\bar{D}_{1\to 2}}{\bar{D}_{3\to\infty}}\varepsilon_{\bar{D}_{1\to 2},b_{3\to\infty}} + \varepsilon_{\bar{S}_3,b_{3\to\infty}} + \varepsilon_{\bar{\bar{D}}_{3\to\infty},b_{3\to\infty}}.$$

Substituting for the aggregate elasticities and re-arranging, we find

$$\begin{aligned}
&\alpha\left[\frac{D^y}{\bar{D}} - \frac{D^y_{3\to\infty}}{\bar{D}_{3\to\infty}}\right]\varepsilon_{D^y,b} + (1-\alpha)\left[\frac{D^z}{\bar{D}} - \frac{D^z_{3\to\infty}}{\bar{D}_{3\to\infty}}\right]\varepsilon_{D^z,b} > \\
&\qquad \alpha\frac{D^y_{1\to 2}}{\bar{D}_{3\to\infty}}\varepsilon_{D^y_{1\to 2},b_{3\to\infty}} + (1-\alpha)\frac{D^z_{1\to 2}}{\bar{D}_{3\to\infty}}\varepsilon_{D^z_{1\to 2},b_{3\to\infty}} + \alpha\frac{D^y_{3\to\infty}}{\bar{D}_{3\to\infty}}\varepsilon_{S_3^y,b_{3\to\infty}} + (1-\alpha)\frac{D^z_{3\to\infty}}{\bar{D}_{3\to\infty}}\varepsilon_{S_3^z,b_{3\to\infty}}.
\end{aligned}$$

Using

$$\alpha\left[\frac{D^y}{\bar{D}} - \frac{D^y_{3\to\infty}}{\bar{D}_{3\to\infty}}\right] + (1-\alpha)\left[\frac{D^z}{\bar{D}} - \frac{D^z_{3\to\infty}}{\bar{D}_{3\to\infty}}\right] = 0,$$

we can re-write the inequality as

$$\begin{aligned}
&\alpha\left[\frac{D^y}{\bar{D}} - \frac{D^y_{3\to\infty}}{\bar{D}_{3\to\infty}}\right]\left[\varepsilon_{D^y,b} - \varepsilon_{D^z,b}\right] > \\
&\qquad \alpha\frac{D^y_{3\to\infty}}{\bar{D}_{3\to\infty}}\left[\frac{D^y_{1\to 2}}{D^y_{3\to\infty}}\varepsilon_{D^y_{1\to 2},b_{3\to\infty}} + \varepsilon_{S_3^y,b_{3\to\infty}}\right] + (1-\alpha)\frac{D^z_{3\to\infty}}{\bar{D}_{3\to\infty}}\left[\frac{D^z_{1\to 2}}{D^z_{3\to\infty}}\varepsilon_{D^z_{1\to 2},b_{3\to\infty}} + \varepsilon_{S_3^z,b_{3\to\infty}}\right].
\end{aligned}$$

At this point we can see the mechanism at work. The LHS of the inequality can be made larger by increasing h_1^y relative to h_1^z. The y-type agents are more responsive to changes in benefits (i.e., $\partial\varepsilon_{D^y,b}/\partial h_1^y > 0$) and spend relatively less time unemployed later in the spell (i.e., $\partial(D^y/D^y_{3\to\infty})/\partial h_1^y > 0$). At the same time, we can make the right-hand side arbitrarily small by increasing the heterogeneity. As we increase h_1^y and decrease h_1^z, the forward looking elasticities of the y-types increase while the same elasticities decrease for the z-types. However, increasingly little weight (converging to zero for $h_0 + h_1^y s_y^\rho \to 1$) gets placed on the y-types' elasticity. More weight gets placed on the forward-looking elasticities of the z-types, but these are low and converge to zero for $h_1^z \to 0$.

Hence, with sufficient heterogeneity, we have that $MH_{1\to\infty} > MH_{3\to\infty}$.

A.4.5 Heterogeneity in Assets

Having introduced heterogeneity in exit rates, it is straightforward to reverse the prediction on the gradient of the consumption smoothing gains as well. This requires individuals with lower marginal utility of consumption to select into longer unemployment spells in a way that the dynamic selection offsets the increase in marginal utility for a given individual due to the depletion of assets. This can be obtained for example by heterogeneity in assets where an agent's asset holdings are negatively correlated with her exit rate. The same argument applies with heterogeneity in preferences.

To illustrate this, consider our two-type setup where type y is borrowing constrained and has high exit rate h^y - potentially induced by the constrained consumption when unemployed - and type z who has access to assets and low exit rate $h^z < h^y$. To obtain $CS_t > CS_{t'}$, we need

$$\alpha \frac{S_t^y}{S_t} u'\left(c_t^y\right) + (1-\alpha) \frac{S_t^z}{S_t} u'\left(c_t^z\right) > \alpha \frac{S_{t+1}^y}{S_{t+1}} u'\left(c_{t+1}^y\right) + (1-\alpha) \frac{S_{t+1}^z}{S_{t+1ds}} u'\left(c_{t+1}^z\right)$$

$\Leftrightarrow$

$$\alpha \left[\frac{S_t^y}{S_t} - \frac{S_{t+1}^y}{S_{t+1}}\right] u'(b) + (1-\alpha) \left[\frac{S_t^z}{S_t} - \frac{S_{t+1}^z}{S_{t+1}}\right] u'\left(c_t^z\right) + (1-\alpha) \frac{S_{t+1}^z}{S_{t+1}} \left[1 - \frac{u'\left(c_{t+1}^z\right)}{u'\left(c_t^z\right)}\right] u'\left(c_t^z\right) > 0$$

$\Leftrightarrow$

$$\alpha \left[\frac{S_t^y}{S_t} - \frac{S_{t+1}^y}{S_{t+1}}\right] \left[u'(b) - u'\left(c_t^z\right)\right] + (1-\alpha) \frac{S_{t+1}^z}{S_{t+1}} \left[u'\left(c_t^z\right) - u'\left(c_{t+1}^z\right)\right] > 0$$

Now notice that with $h_1^y > h_1^z$, the relative survival rate of agents of type y is decreasing over the spell,

$$\begin{aligned} \frac{S_t^y}{S_t} - \frac{S_{t+1}^y}{S_{t+1}} &= \frac{(1-h^y)^{t-1}}{\alpha (1-h^y)^{t-1} + (1-\alpha)(1-h^z)^{t-1}} - \frac{(1-h^y)^t}{\alpha (1-h^y)^t + (1-\alpha)(1-h^z)^t} \\ &= \frac{\left[\frac{S_{t+1}}{S_t} - (1-h^y)\right](1-h^y)^{t-1}}{\alpha (1-h^y)^t + (1-\alpha)(1-h^z)^t} > 0 \end{aligned}$$

since $S_{t+1}/S_t > (1-h^y)$. Moreover, the difference in marginal utility of consumption is positive, $u'(b) - u'\left(c_t^z\right) > 0$, and more so the higher the asset level of agents of type z. Hence, with sufficient heterogeneity in the exit rates and the asset levels, the selection effect can offset the increase in marginal utility for agents of type z, $u'\left(c_t^z\right) - u'\left(c_{t+1}^z\right) < 0$, and as such make the consumption smoothing gains higher for benefits timed earlier in the spell.

A.4.6 Relative Survival Rate Response

We use the model with heterogeneity in search efficacy to show that the relative survival rate in a two-type model can increase in response to benefits paid early in the spell and decrease in response to benefits paid later in the spell. Embedding this in a model with employer screening considered in would imply that the gradient of the moral hazard cost could become steeper when adjusted for the hiring externality.

In particular, we are interested in

$$\frac{\partial h_{t'}}{\partial b_t} / S_t \propto \frac{\partial \left[S_{t'}^H / S_{t'}^L\right]}{\partial b_t} / S_t = S_{t'}^H / S_{t'}^L \frac{\frac{\partial S_{t'}^H}{\partial b_t} / S_{t'}^H - \frac{\partial S_{t'}^L}{\partial b_t} / S_{t'}^H}{S_t}.$$

We now analyze how $\frac{\partial S_{t'}^i}{\partial b_t} / S_{t'}^i$ changes when increasing search efficacy $h_{1,i}$. For simplicity, we assume $h_0 = 0$ and denote $h_{1,i} = h_i$. Using similar steps as before and starting again from a flat profile, we find

$$\frac{\partial S_{t'}^i}{\partial b_t} / S_t^i = \frac{h_i s_i^\rho}{\beta} \sum_{j=0}^{t'} [\beta(1 - h_i s_i^\rho)]^{t-j-2} \frac{B}{b}$$

for $t > t'$. Like before, assuming that $\partial B / \partial h_i \approx 0$ and $\partial h_i s_i^\rho / \partial h_i \approx s_i^\rho$, which follows from $1 - \beta \approx 0$, we find

$$\frac{\partial}{\partial h_i} \left[\frac{\partial S_{t'}^i}{\partial b_t} / S_t^i\right] \approx \left[s^\rho \sum_{j=0}^{t'} [\beta(1 - hs^\rho)]^{t-j-2} - s^\rho \frac{hs^\rho}{1 - hs^\rho} \sum_{j=0}^{t'} (t-j-2)[\beta(1 - hs^\rho)]^{t-j-2}\right] B/b$$

We now want to see if this term can be negative for high enough t, but positive for low enough t. First,

$$\frac{\partial}{\partial h_i}\left[\frac{\partial S^i_{t'}}{\partial b_t}/S^i_t\right] < \left[x^\rho \sum_{j=0}^{t'}[\beta(1-hs^\rho)]^{t-j-2} - x^\rho \frac{hs^\rho}{1-hs^\rho}(t-t'-2)\sum_{j=0}^{t'}[\beta(1-hs^\rho)]^{t-j-2}\right]B/b$$
$$= 1 - \frac{hs^\rho}{1-hs^\rho}(t-t'-2) < 0.$$

The last inequality holds for high enough t (provided $hs^\rho > 0$). Second,

$$\frac{\partial}{\partial h_i}\left[\frac{\partial S^i_{t'}}{\partial b_t}/S^i_t\right] > \left[s^\rho \sum_{j=0}^{t'}[\beta(1-hs^\rho)]^{t-j-2} - s^\rho \frac{hs^\rho}{1-hs^\rho}(t-2)\sum_{j=0}^{t'}[\beta(1-hs^\rho)]^{t-j-2}\right]B/b$$
$$= 1 - \frac{hs^\rho}{1-hs^\rho}(t-2) > 0.$$

The last inequality now holds for low enough t (provided hs^ρ is small).

FOR ONLINE PUBLICATION - Appendix B: Additional results and robustness of the RK design

This Appendix presents additional results on the duration responses to benefits and various robustness checks of the RK design.

B.1 Additional Results: Hazard Rate Responses

To further investigate the non-stationary patterns in unemployment responses, Figure B-1 reports the RKD estimates of the effect of UI benefits on the hazard rates out of unemployment.

Since hazard rates are quite noisy at very high frequency, we have defined hazard rates by periods of 5 weeks. Blue dots represent the marginal effect of a change in both b_1 and b_2, estimated in the regression kink design for spells starting between 1999 and July 2001. Red dots represent the marginal effect of a change in b_2 only, estimated in the regression kink design for spells starting between July 2001 and July 2002. In both cases, 95% confidence interval around the point estimates, from robust standard errors, are displayed. The figure conveys quite clearly a series of interesting findings.

First, the graph shows that the effect of UI benefits is mostly concentrated in the first 10 to 15 weeks. After 15 weeks, the effect of UI benefits on the hazard rate is small and almost always insignificant.

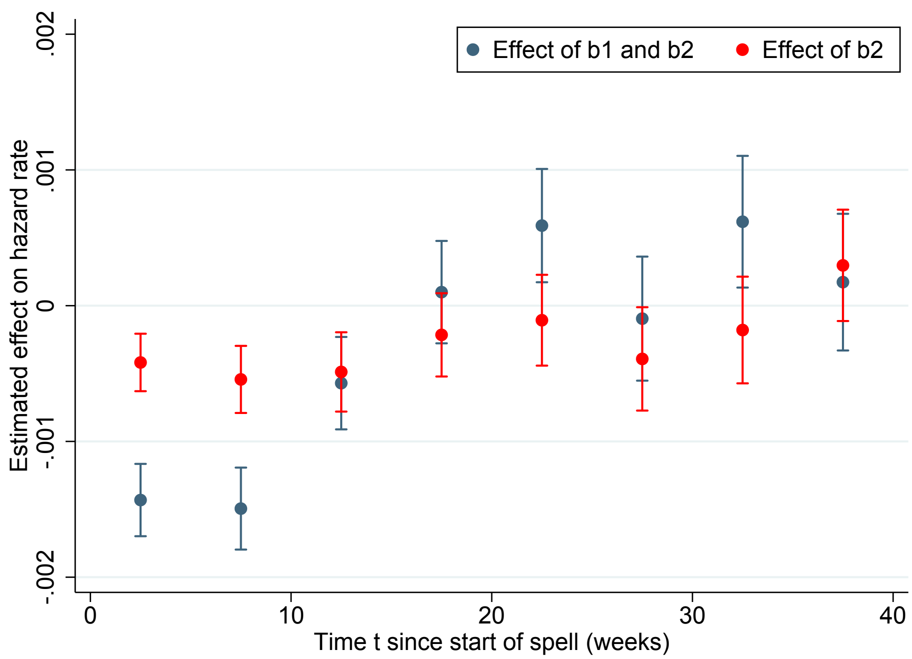
Second, the graph shows that b_2 (benefits received after 20 weeks) do have an effect on the hazard rate in the first 10 weeks. This confirms that unemployed individuals are forward-looking. b_2 does have a somewhat negative effect on contemporaneous hazard rates (after 20 weeks), but this effect is small and almost always insignificant.

The effect of b_1 can easily be inferred as it is the difference, for each hazard rate, between the effect of b_1 and b_2, and the effect of b_2 only. From the figure, we can easily see that the effect of b_1 is almost twice as large as the effect of b_2 early on in the spell. Because hazard rates are very responsive to b_1 in the spell, b_1 is having a large effect on the probability to survive into unemployment after 20 weeks. This creates a large mechanical effect of b_1 on D_2, the average duration spent in the second part of the benefit profile.

The total effect of b_1 on D_2 is the sum of the mechanical effect on survival plus the effect of b_1 on hazard rates after 20 weeks. Interestingly, the figure shows that the latter effect is positive (though small) for some hazard rates after 20 weeks. This is an indication of some (positive) dynamic selection going on: individuals who remain unemployed due to higher b_1 have a slightly higher hazard rate later in the spell. Yet, this dynamic selection effect is not large enough to undo the large mechanical effect that a much larger fraction of individuals survive into the second part of the benefit profile.

The figure therefore provides some intuition for why b_1 has a MH cost that is somewhat larger than b_2. b_1 increases D_1 more than b_2 because it strongly affects hazard rates early in the spell. This in turn has a large mechanical effect on D_2 since more individuals survive into the second part of the benefit profile. The effects of b_1 (positive) and b_2 (negative) on hazard rates after 20 weeks are too small and insignificant to undo, in the MH costs estimates, the effects on hazard rates early in the spell.

Figure B-1: RKD estimates on hazard rates at the SEK725 kink



Notes: The figure reports the RKD estimates of the effect of UI benefits on the hazard rates out of unemployment. Empirical hazard rates are the observed fraction of individuals exiting unemployment in period t conditional on surviving until the start of period t, and are defined by periods of 5 weeks. Blue dots represent the marginal effect of a change in both $b_{\mathbf{1}}$ and $b_{\mathbf{2}}$, estimated in the regression kink design for spells starting between 1999 and July 2001. Red dots represent the marginal effect of a change in $b_{\mathbf{2}}$ only, estimated in the regression kink design for spells starting between July 2001 and July 2002. All estimates are from linear specifications using the changes in the UI schedule at the 725SEK kink with a 90SEK bandwidth. 95% confidence intervals around the point estimates, from robust standard errors, are displayed.

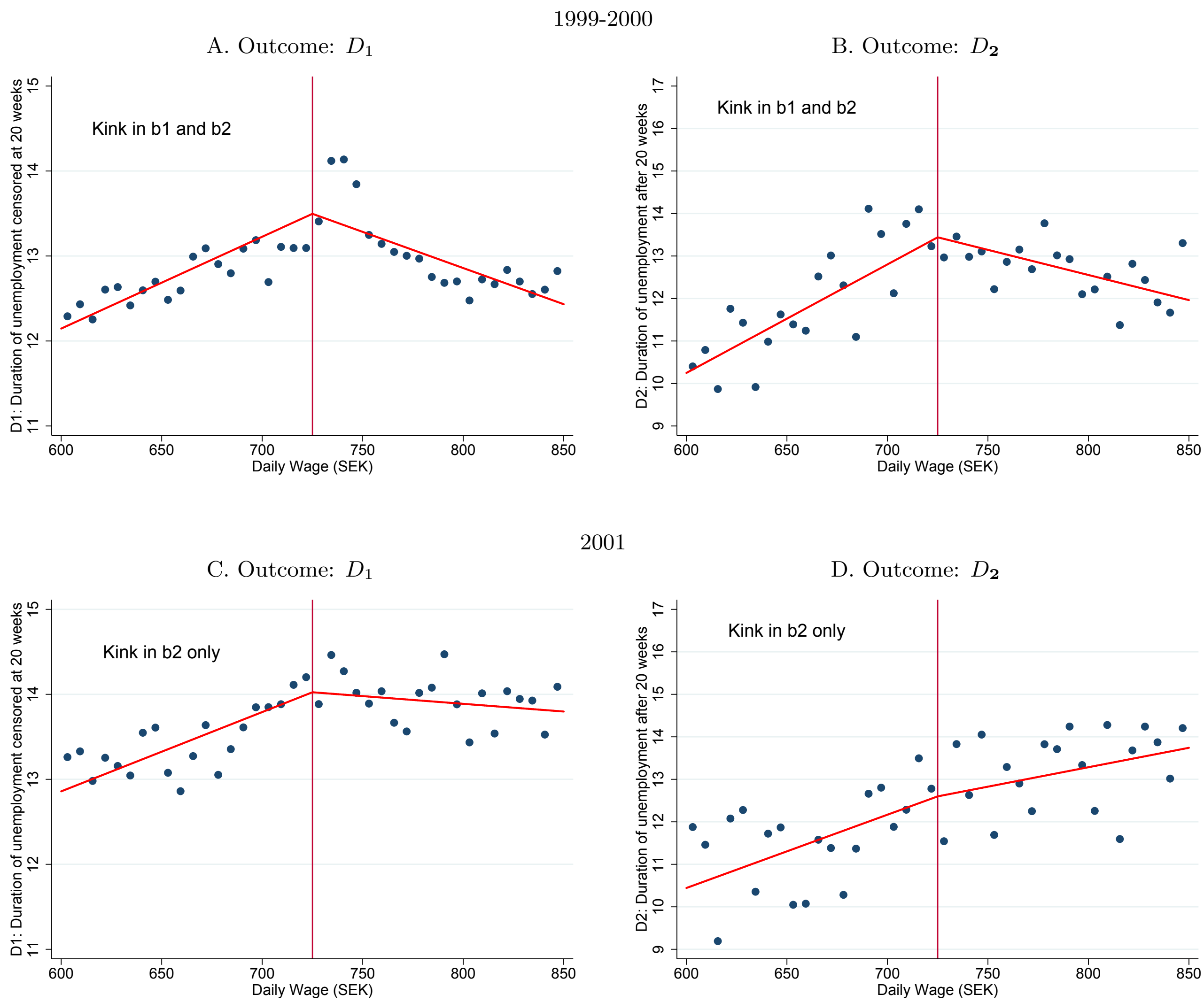
B.2 RK design for D_1 and D_2

To assess the validity of the RK design for unemployment duration D_1 spent on the first part of benefit profile and unemployment duration D_2 spent in the second part of the benefit profile, Figure B-2 below displays the raw data, replicating for D_1 and D_2 what Figure 2 was doing for total unemployment duration D. The graphs provide graphical evidence of a change in slope in the relationship between both D_1 and D_2 and previous daily wage in response to the kink in UI benefits. The change in slope is larger for spells starting before July 2001, when both b_1 and b_2 are capped at the 725SEK threshold. The magnitude of the change in slope decreases for spells starting between July 2001 and July 2002 when only b_2 is capped at the 725SEK threshold. Formal estimates of the change in slope using linear specifications of the form of equation (16) are displayed in Table 2. The red lines display predicted values of the regressions in the linear case.

B.3 Year by year RKD estimates

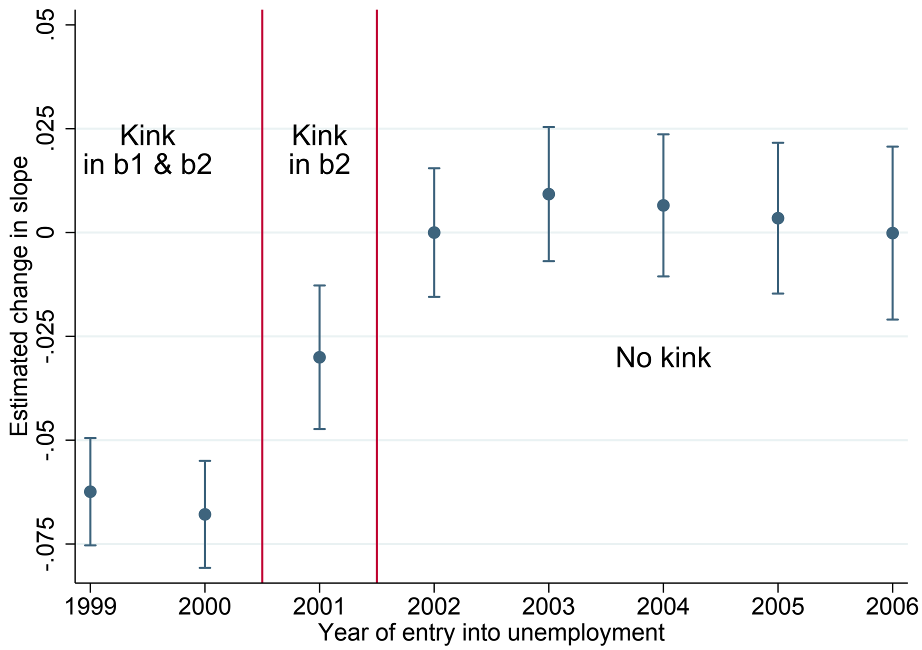
Figure B-3 plots the year-by-year evolution of the estimates of the change in slope in the relationship between total unemployment duration D and pre-unemployment daily wages from 1999 to 2007. The figure provides clear evidence that our estimated responses are indeed due to the policy changes, and not due to time trends in the distribution of durations around the kink.

Figure B-2: RK design at the SEK725 threshold for D_1 and D_2



Notes: The Figure plots average unemployment duration D_1 spent on the first part of benefit profile and average unemployment duration D_2 spent on the second part of the benefit profile, in bins of previous daily wage for the two periods of interest. Sample is restricted to unemployed individuals with no earnings who report being searching for full-time employment. The graphs provide graphical evidence of a change in slope in the relationship between both D_1 and D_2 and previous daily wage in response to the kink in UI benefits. The change in slope is larger for spells starting before July 2001, when both b_1 and b_2 are capped at the 725SEK threshold. The magnitude of the change in slope decreases for spells starting between July 2001 and July 2002 when only b_2 is capped at the 725SEK threshold. Formal estimates of the change in slope using linear specifications of the form of equation (16) are displayed in Table 2. The red lines display predicted values of the regressions in the linear case.

Figure B-3: RKD estimates on unemployment duration D at the SEK725 kink by year of entry



Notes: The figure reports the RKD estimates of the effect of UI benefits on total duration of unemployment by year of entry into unemployment, at the 725SEK kink. Entry into unemployment in Year N is defined as starting a spell between of July 1st of Year N and July 1st of Year $N+1$. Spells starting before 2001 are therefore subject to a kink in both b_1 and b_2. Spells starting in 2001 are subject to a kink in b_2 only. Spells starting in 2002 and after do not face any kink in the schedule and represent a placebo. All estimates are from linear specifications using the changes in the UI schedule at the 725SEK kink with a 90SEK bandwidth. 95% confidence intervals around the point estimates, from robust standard errors, are displayed. The figure provides clear evidence that estimated responses in the RK design are indeed due to the policy changes, and not due to time trends in the distribution of durations around the kink.

B.4 Additional robustness analysis of the RK design

This subsection presents various additional robustness checks of the RK design. We start by restating the two fundamental identifying assumptions of the RK design, and then propose various tests to assess their potential validity, by looking for clear violations of these assumptions.

We consider the general model:

$$Y = y(b_{\mathbf{1}}, b_{\mathbf{2}}, w, \mu),$$

We are interested in identifying the marginal effect of benefits $b_k, k = 1, 2$ on the duration outcome Y, $\alpha_{\mathbf{k}} = \frac{\partial Y}{\partial b_{\mathbf{k}}}$. $b_{\mathbf{k}}$ is a deterministic, continuous function of the wage w, kinked at $w = \bar{w}_{\mathbf{k}}$. Identification of $\alpha_{\mathbf{k}}$ in the RK design relies on two assumptions:

Assumption 1: the direct marginal effect of the assignment variable w on Y is assumed to be smooth around the kink point $\bar{w}_{\mathbf{k}}$. This means that $\frac{\partial y(b_{\mathbf{1}}, b_{\mathbf{2}}, w, \mu)}{\partial w}$ is assumed to be continuous in the neighborhood of the kink point.

Assumption 2: the distribution of unobserved heterogeneity μ is assumed to be evolving smoothly around the kink point. This means that the conditional density $(f_{w|\mu}(\cdot))$ and its partial derivative with respect to w, $(\partial f_{w|\mu}(\cdot)/\partial w)$ are assumed to be continuous in the neighbourhood of the kink point.

These identifying assumptions are, by definition, untestable. Yet, we can use the various “experiment arms” of our quasi-experimental setting as well as sensitivity analysis to try to detect clear violations of these assumptions and to provide some sense of the potential robustness of these identifying assumptions and the validity of our RK design.

Testing for clear violations of Assumption 2: manipulation The most obvious violation of the assumption of smooth distribution of heterogeneity at the kink arises if individuals are able to locate their daily wage strategically around the kink point. A few tests can help assess the robustness of this assumption.

First, Figure B-4 plots the density of the daily wage and shows graphically the smoothness of the distribution of the assignment variable at the kink point in the UI schedules. The graph shows the probability density function of the daily wage around the 725SEK threshold and displays two formal tests. The first is a standard McCrary test of the discontinuity of the pdf of the assignment variable. We report the difference in height of the pdf at the threshold. The second is a test for the continuity of the first derivative of the p.d.f. We report the coefficient estimate of the change in slope of the pdf in a regression of the number of individuals in each bin on polynomials of the assignment variable interacted with a dummy for being above the threshold. Both tests suggest smoothness of the assignment variable around the threshold

The continuity in the pdf of the assignment variable indicates that there is no bunching at the kink point. Such bunching would have constituted proof of the ability of individuals to manipulate their location on the UI schedule, which would have been a clear violation of Assumption 2. Absence of bunching at the kink is not a sufficient condition to rule out that individuals respond to the kinked schedule in their earnings decision, which would question the validity of Assumption 2. The absence of bunching could be driven by optimization frictions which attenuate the ability to bunch at the kink, or by the fact that the compensated elasticity of daily wage with respect to marginal tax rates is small. Even if the compensated elasticity of the daily wage is small, income effects could still be large, and would affect earnings decisions as we move further away from the kink. This would then be picked up by variations in the slope of the pdf at the kink. The fact that we do not detect any change in the first derivative of the pdf of daily wage at the kink point, as reported in Figure B-4, is reassuring.

Interestingly, because the kinks in the schedule of $b_{\mathbf{1}}$ and $b_{\mathbf{2}}$ are removed in July 2001 and July 2002, we can actually directly estimate whether the distribution of daily wages reacts to the removal of the kink and therefore get a direct test of whether the pdf of the assignment variable is affected by the presence of the kink. In Table B-1 below, we report the results of a difference-in-difference model where we look at the evolution of log wages above and below the kink, before July 2001 (when both kinks were in place) and after July 2001 (when one kink is removed).

The wages of individuals who had optimally chosen their daily wages at or above the kink, will be affected by the removal of the kink. To the contrary, individuals who had optimally chosen daily wages below the kink should not be affected by the removal of the kink. If individuals' daily wages respond to the kinked UI schedule, we therefore expect a differential change in the average log wages above the kink after July 2001 relative to log wages below the kink. Estimates, reported in Table B-1 indicate that the removal of the kinks did not significantly affect the distribution of daily wages above and below the kink. There is no differential change in the daily wage below and above the kink after July 2001. This in turn suggests that the presence of kinks in the UI schedule does not significantly affect the distribution of daily wages around the kink.

Testing for clear violations of Assumption 2: observable heterogeneity To further investigate the evolution of the distribution of heterogeneity at the kink, the panels in Figure B-5 show how the mean values of different covariates (age, fraction of men, highly educated and foreigners) evolve with the daily wage around the kink. We do not find any non-linearity around the kink. This is also reassuring, as non-smoothness in the distribution of observable heterogeneity would have cast doubt on the validity of the assumption of smoothness in the distribution of unobservable heterogeneity around the kink.

Testing for underlying non-linearities: Bandwidth size The panels in Figure B-6 report our RKD estimates for different bandwidth sizes. For all periods we consider, the estimates remain stable for bandwidths above $h =$ 60SEK.

Testing for underlying non-linearities: Permutation tests Ganong and Jaeger [2014] suggest that it can be helpful to assess whether the true coefficient estimate is larger than those at "placebo" kinks placed away from the true kink. The idea behind their permutation test is that, if the counter-factual relationship between the assignment variable and the outcome (i.e., in the absence of the kink in the budget set) is non-linear, then the curvature in this relationship will result in many of the placebo estimates being large and statistically significant. In Table B-2, we report 95% confidence interval based on this permutation procedure and compare them to bootstrapped standard errors and robust standard errors.

Testing for underlying non-linearities: Non-parametric detection of kink point Figure B-10 shows the R-squared when we run the RKD regression in (16) for "placebo" kinks placed in 10SEK increments from the true location of the threshold. This procedure, proposed in Landais [2015], and inspired from the time series literature on detection of trend breaks, enables to non-parametrically detect where a true kink is most likely to be located in the data, by looking at the placebo kink where the R-squared is maximized. The figure shows that the R-squared is maximized at the location of the actual kink point, again supporting the evidence that there is in fact a change in slope that occurs at the actual kink point. In both panels A and B, the preferred location of the kink is extremely close to the true kink and the relationship between the placebo kink location and R-squared of the model exhibits a clear concave shape. In panel C, reassuringly, when there is no true kink at 725SEK, this relationship is perfectly flat.

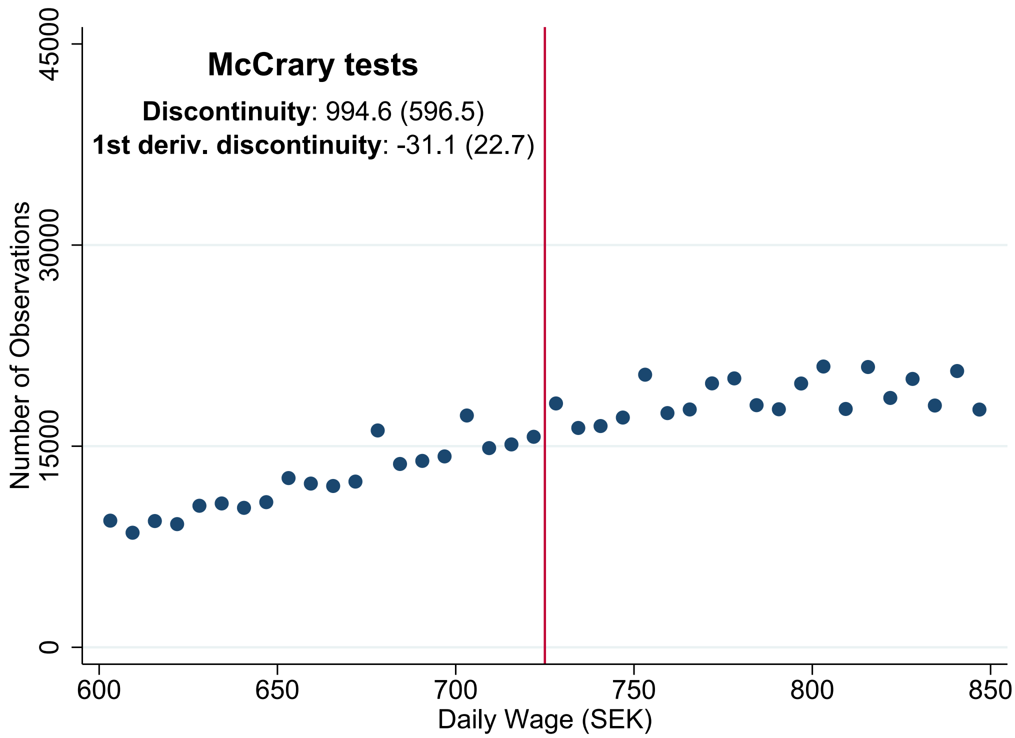
Polynomial order Table B-3 shows estimates of the change in slope at the kink for linear, quadratic and cubic specifications, assessing the model fit for these different specifications.

For the 1999-2000 period, the estimates are very similar across polynomial orders. For the 2001 period, estimates vary across polynomial orders, and estimates from the quadratic model are larger in magnitude than estimates using a linear specification. Yet, model fit analysis suggests that linear estimates should be preferred. The linear specification is having similar root mean squared errors (RMSE) and minimizes the Aikake information criterion (AIC). Note also that, although larger, the point estimates on the quadratic specification are very imprecisely estimated, so that we cannot actually reject that they are equal to the estimates from the linear model.

We also plot below in Figure B-8 the prediction from the linear and quadratic specifications on top of the raw data to see how these models fit the data. For the period 1999-2000, panel A shows that both the quadratic and the linear model fit the data equally well and deliver extremely similar results for the change in slope at the kink. For the period 2001, the quadratic model delivers a larger change in slope at the kink compared to the linear fit. But this is driven by a higher curvature so that the linear model overall does deliver a better fit of the data, as indicated by the root squared mean error and the AIC reported on the graph.

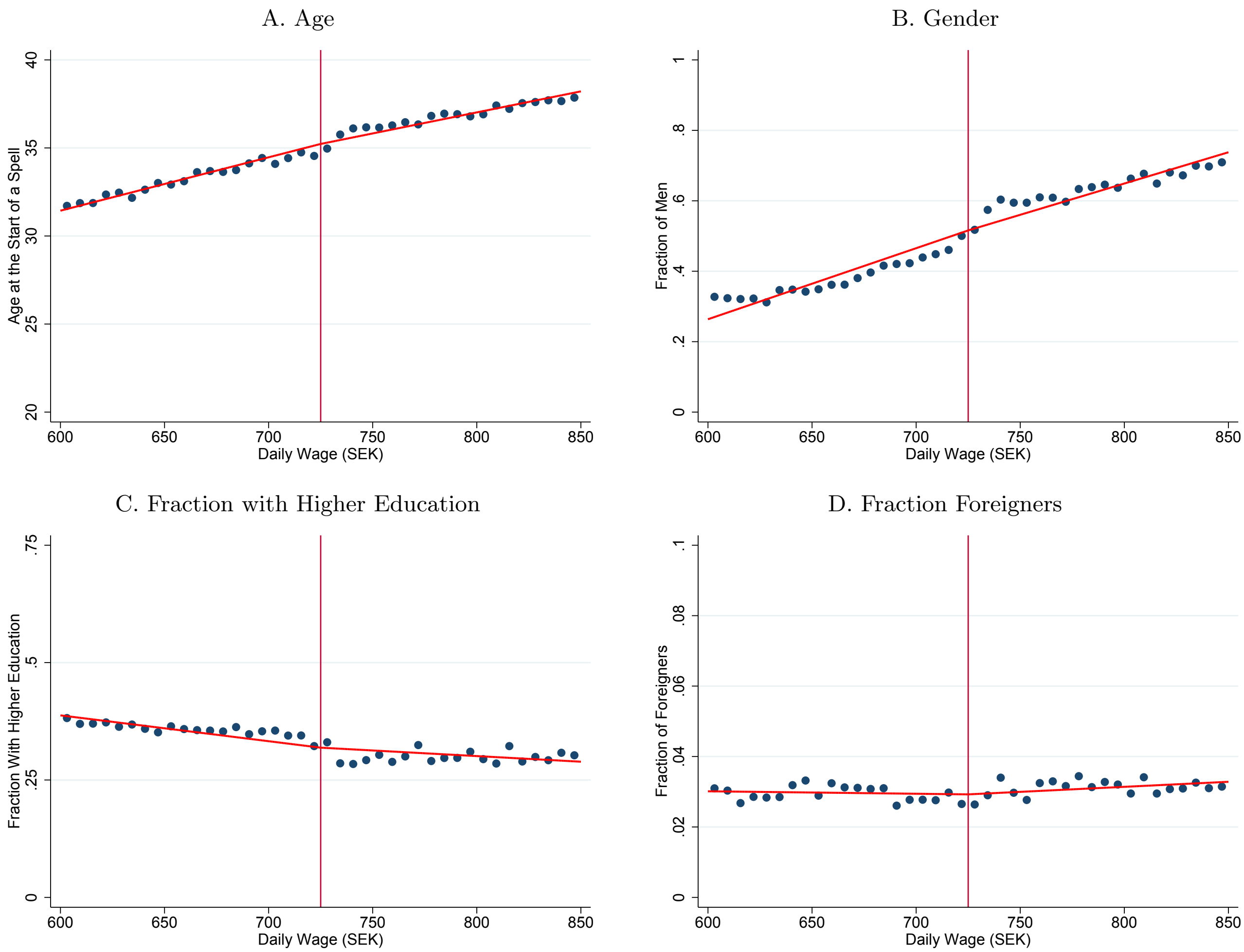
Right-censoring When the schedule of benefits changes, individuals with ongoing spells are transferred to the new schedule. To control for this, two solutions can be envisaged. First, one may get rid of observations who have an ongoing spell at the moment the schedule changes. An alternative solution is to treat the duration of these observations as censored at the moment when these individuals transfer the new schedule. One can then estimate a Tobit model on the right-censored data. In Figure B-9 below, we report the estimates for the estimated change in slope in $D_{\mathbf{1}}$ and $D_{\mathbf{2}}$ for censored and uncensored models, as a function of the RKD bandwidth. The Figure shows that censored and uncensored models deliver identical results, and that the point estimates of the two models are never statistically significantly different. The uncensored model proves a little less precise though, as we end up throwing away some observations. As a consequence, we have decided to focus on the estimates from these censored models for our baseline results.

Figure B-4: Robustness of the RK design: p.d.f of daily wage



Notes: The figure tests graphically the smoothness of the distribution of the assignment variable at the kink point in the UI schedules to assess the validity of the local random assignment assumption underlying the RK design. The Panel shows the probability density function of the daily wage around the 725SEK threshold. We also display two formal tests of the identifying assumptions of the RKD. The first is a standard McCrary test of the discontinuity of the p.d.f of the assignment variable. We report the difference in height of the p.d.f at the threshold. The second is a test for the continuity of the first derivative of the p.d.f. We report the coefficient estimate of the change in slope of the p.d.f in a regression of the number of individuals in each bin on polynomials of the assignment variable interacted with a dummy for being above the threshold. Both tests suggest smoothness of the assignment variable around the threshold, in support of the identifying assumptions of the RK design.

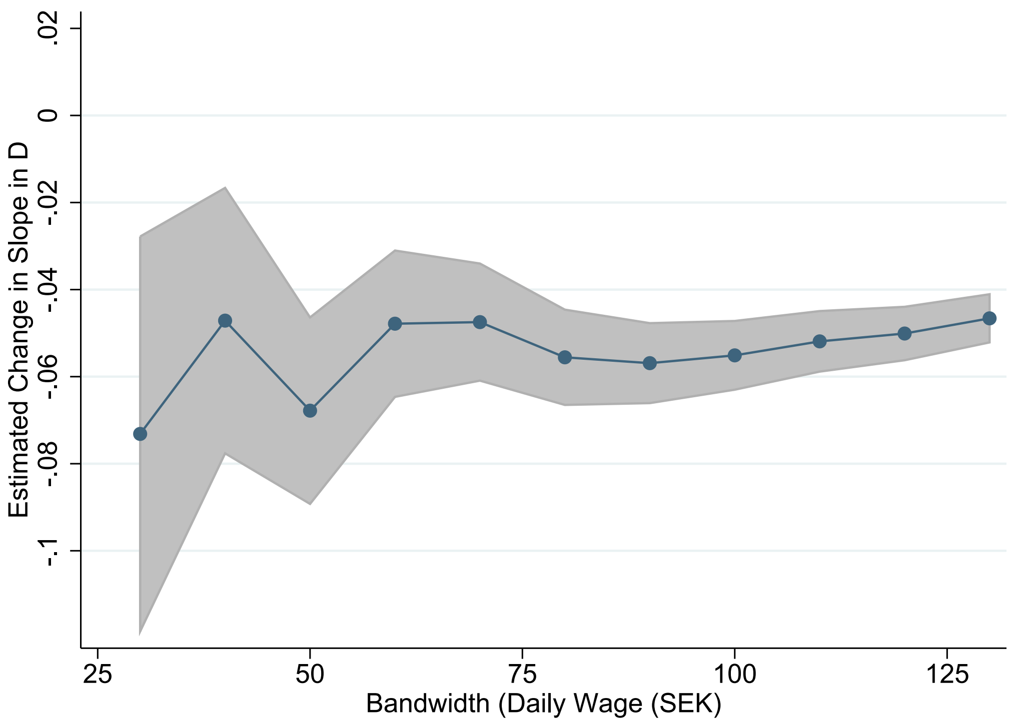
Figure B-5: Robustness of the RK design: Covariates



Notes: The figure tests the validity of the smoothness assumptions of the RK design. Each panel shows the mean values of a different covariate in bins of the assignment variable around the 725SEK threshold. The red lines display predicted values of polynomial regressions of the form of equation (16) in order to detect potential non-linearity around the threshold. The sample is restricted to all spells starting before July 2002, when kinks in the UI schedule are active at the 725SEK threshold. The graphs show evidence of smoothness in the evolution of all covariates at the kink, in support of the RKD identification assumptions.

Figure B-6: RKD estimates as a function of bandwidth size

A. **1999 - 2000**



B. **2001**

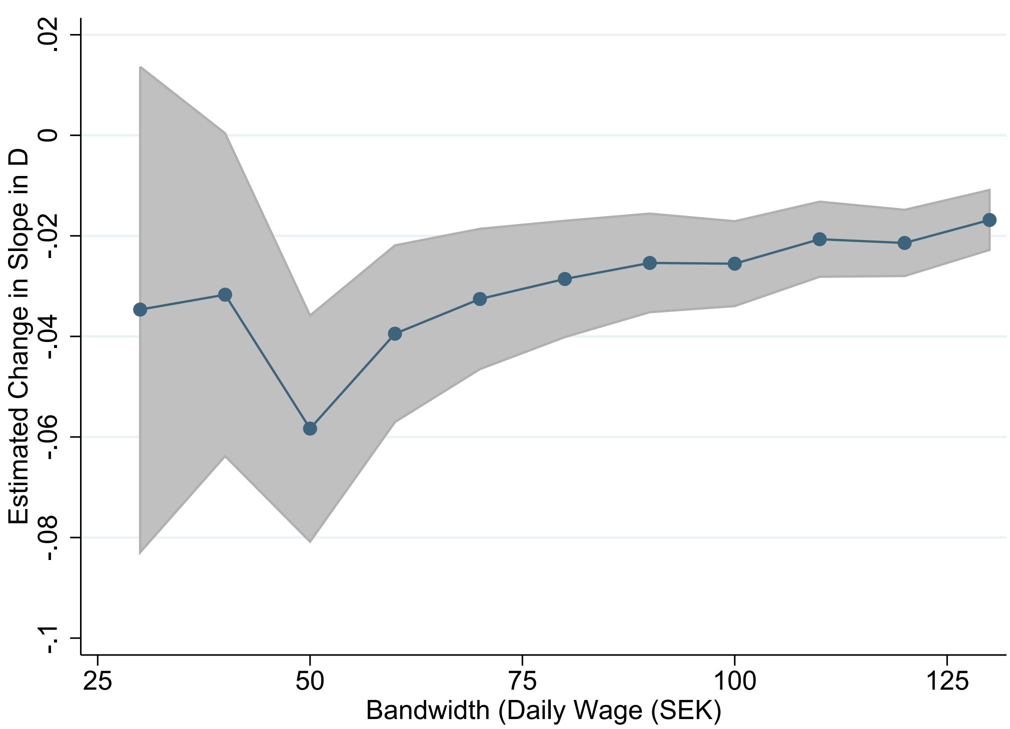
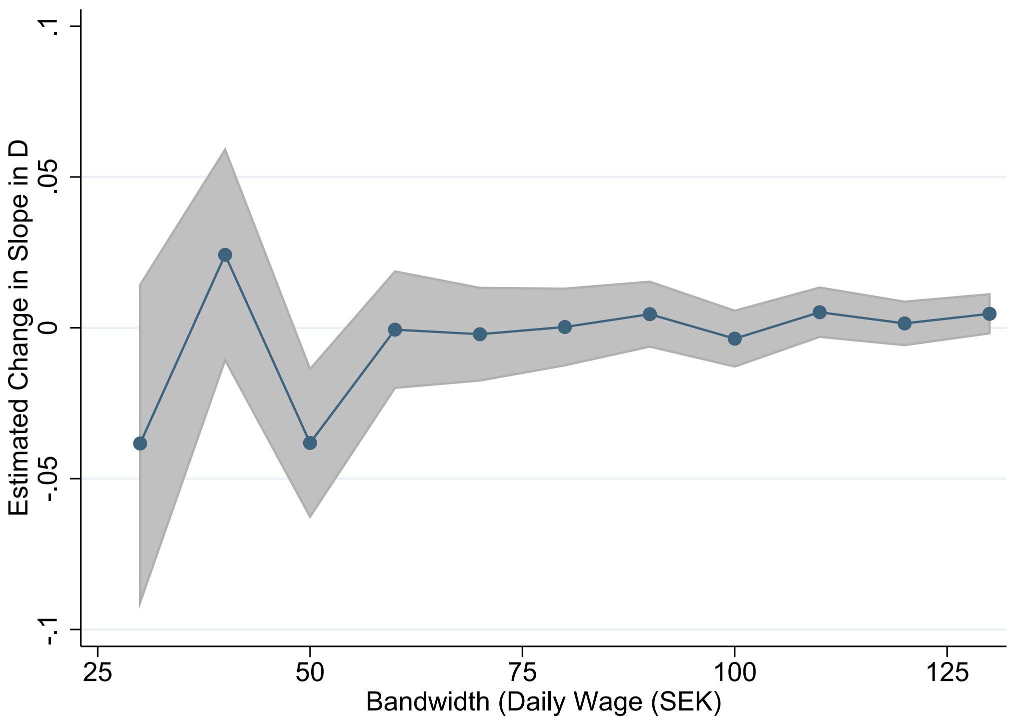


Figure B-7: RKD ESTIMATES AS A FUNCTION OF BANDWIDTH SIZE (*continued*)

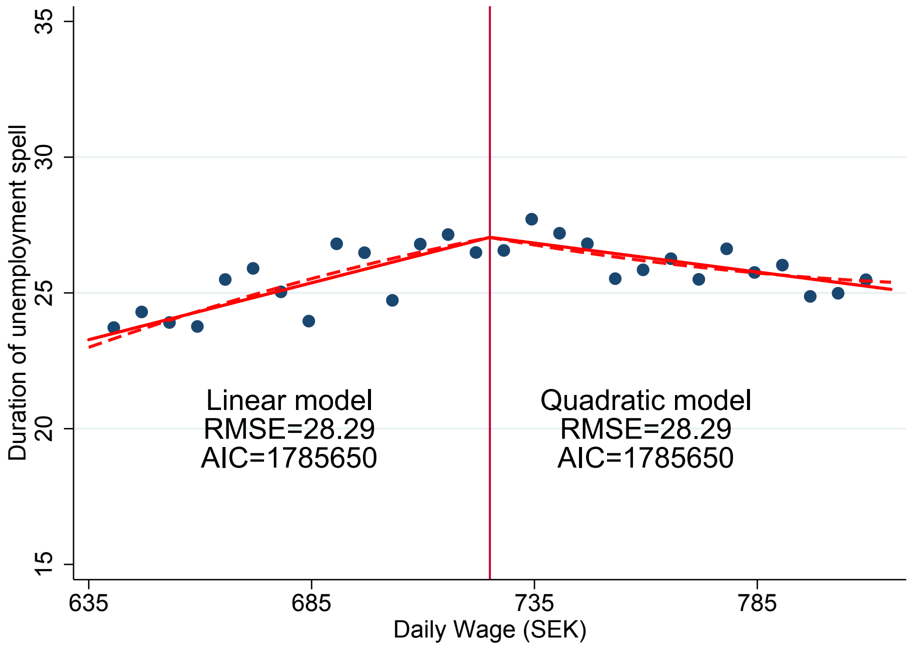
A. **2002-2005**



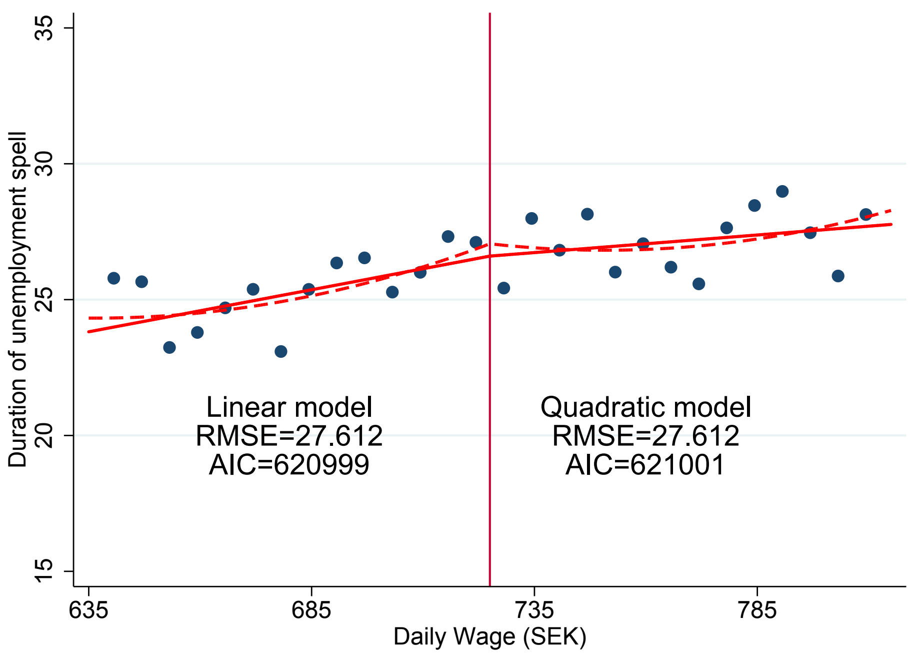
Notes: The figure reports estimates of the change in slope with 95% robust confidence interval in the relationship between unemployment duration and daily wage at the 725SEK threshold using linear regressions of the form of equation (16) as a function of bandwidth size h. These estimates are reported for three periods of interest: 1999-2000 (i.e., spells starting before July 2001), 2001 (i.e., spells starting after July 2001 and before July 2002) and 2002- (i.e., spells starting after July 2002). Unemployment duration is defined as the number of weeks between registration at the PES and exiting the PES or finding any employment (part-time or full-time employment, entering a PES program with subsidized work or training, etc.). Unemployment duration is capped at two years. Sample is restricted to unemployed individuals with no earnings who report being searching for full-time employment.

Figure B-8: Unemployment duration as a function of daily wage around the 725SEK kink, and linear and quadratic model fits

A. 1999-2000

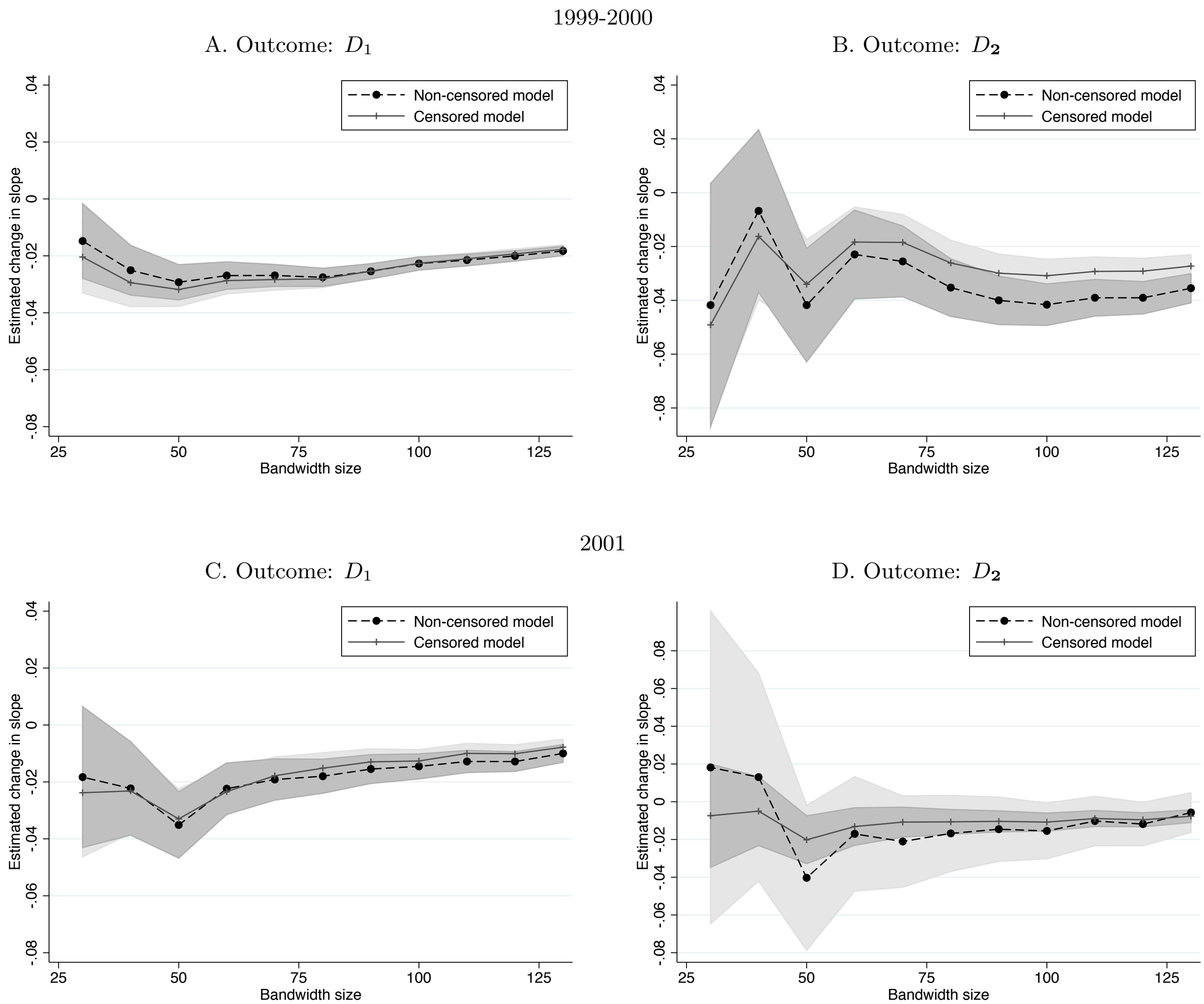


B. 2001



Notes: The figure plots average unemployment duration in bins of previous daily wage for spells starting before July 2001 (panel A) and for spells starting between July 2001 and July 2002 (panel B). On top of the raw data, the figure also displays predictions from linear and quadratic regressions of the form of equation (16) with a bandwidth size $h = 90$SEK. To further assess model fit, we report for each specification the root mean squared error (RMSE) as well as the Aikake information criterion (AIC).

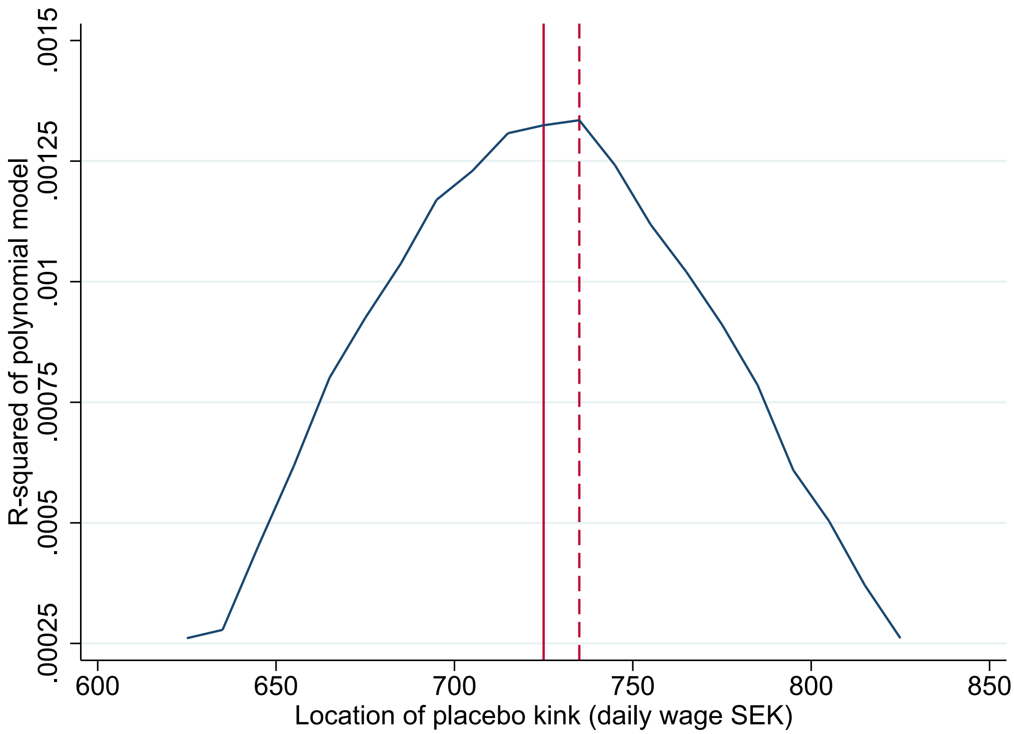
Figure B-9: RKD estimates of the change in slope at the SEK725 kink for OLS model and for the censored regression model



Notes: The figure reports estimates of the change in slope with 95% robust confidence interval in the relationship between unemployment duration and daily wage at the 725SEK threshold using linear regressions of the form of equation (16) as a function of bandwidth size h. When the schedule of benefits changes, individuals who have ongoing spells are transferred to the new schedule. The Figure compares results for two different solutions to account for this. First, one may estimate OLS regressions on a sample where observations who have an ongoing spell at the moment the schedule changes are thrown out (non-censored model). An alternative solution is to treat the duration of these observations as censored at the point when these individuals get in the new schedule. One can then estimate a Tobit model on the right-censored data (censored model). The Figure compares estimates from these two solutions. These estimates are reported for two periods of interest: 1999-2000 (i.e., spells starting before July 2001) and 2001 (i.e., spells starting after July 2001 and before July 2002).

Figure B-10: Non-parametric detection of kink location

A. 1999 - 2000



B. 2001

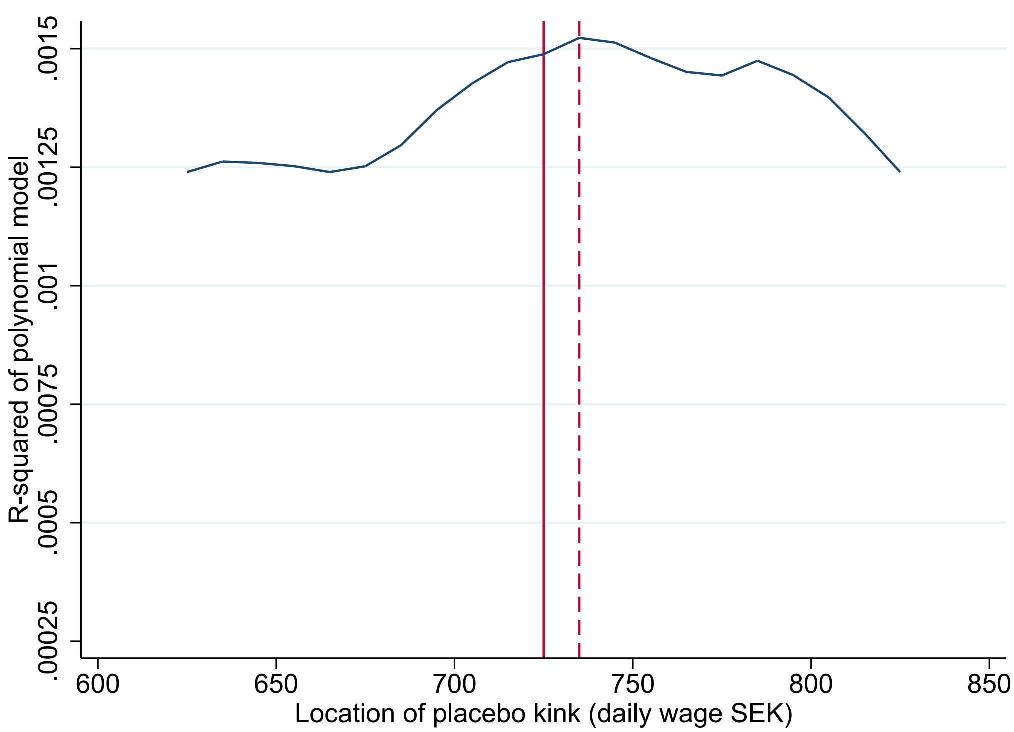
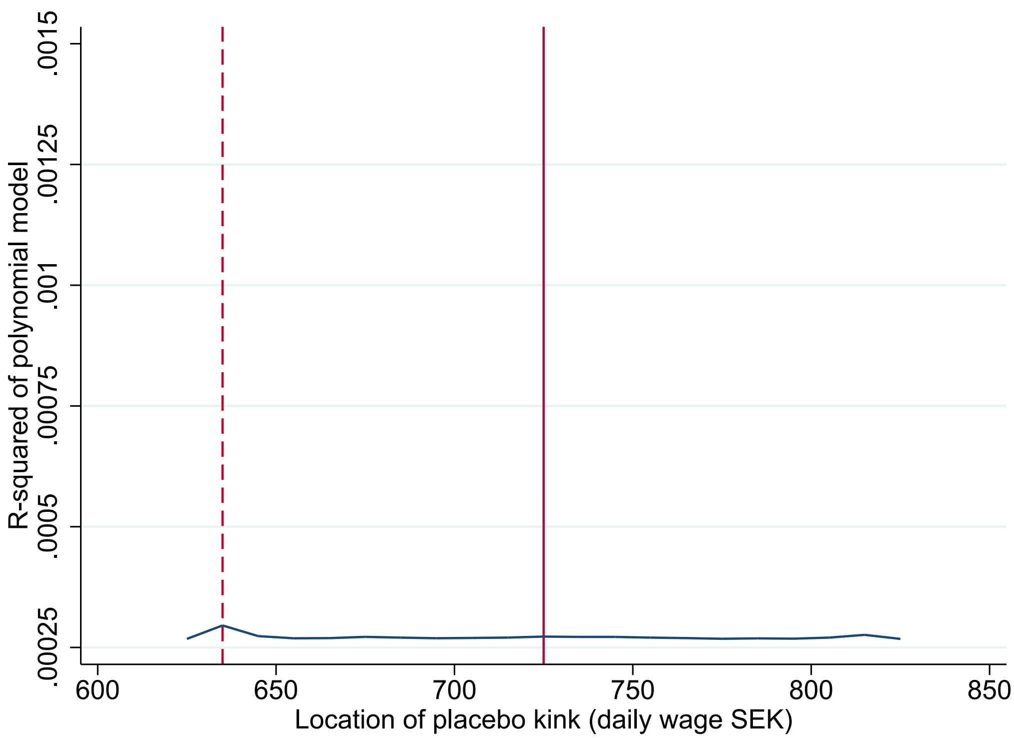


Figure B-11: Non-parametric detection of kink location (*continued)*

C. 2002-2005



Notes: The figure reports the R-squared of polynomial regressions of the form of equation (16) for alternative (placebo) locations of the kink point $\bar{w}_k$ for all observations with wages between 625SEK and 825SEK. The red line indicates the location of the true kink in the schedule. The dashed red line indicates the preferred location of the kink non-parametrically detected in the data, maximizing the R-squared of the model. These estimates are reported for three periods of interest: 1999-2000 (i.e., spells starting before July 2001), 2001 (i.e., spells starting after July 2001 and before July 2002) and 2002- (i.e., spells starting after July 2002). In both panels A and B, the preferred location of the kink is extremely close to the true kink and the relationship between the placebo kink location and R-squared of the model exhibits a clear concave shape. In panel C, when there is no true kink at 725SEK, this relationship is perfectly flat.

Table B-1: Evolution of daily wages below and above the 725SEK point, before and after kinks in the UI schedule are removed

	(1)	(2)	(3)
Dependent variable:		log of daily wage	
$\mathbb{1}[w > \bar{w}]$	0.120^{***}	0.119^{***}	0.119^{***}
	(0.000146)	(0.000146)	(0.000147)
$\mathbb{1}[\text{Spell} > \text{July2001}]$	0.00402^{***}	0.00438^{***}	0.00465^{***}
	(0.000164)	(0.000164)	(0.000165)
$\mathbb{1}[\text{Spell} > \text{July2001}] \times \mathbb{1}[w > \bar{w}]$	0.000305	0.000393	0.000519^{*}
	(0.000222)	(0.000222)	(0.000222)
Age F-E		×	×
Education F-E			×
Gender F-E			×
Industry F-E			×
N	424309	424309	424309

Notes: Standard errors in parentheses.
The Table tests for changes in the distribution of daily wages above and below the kink, as kinks in the schedule of b_1 and b_2 are removed in July 2001 and July 2002. The table reports the results of a difference-in-difference model of the form:

$$\log w = \beta_0 + \beta_1 \mathbb{1}[w > \bar{w}] + \beta_2 \mathbb{1}[\text{Spell} > \text{July2001}] + \beta_3 \mathbb{1}[\text{Spell} > \text{July2001}] \times \mathbb{1}[w > \bar{w}] + \eta$$

The wages of individuals who had optimally chosen their daily wages at or above the kink, will be affected by the removal of the kink. To the contrary, individuals who had optimally chosen daily wages below the kink should not be affected by the removal of the kink. If individuals' daily wages respond to the kinked UI schedule, we therefore expect a differential change in the average log wages above the kink after July 2001 relative to log wages below the kink, captured by β_3. Estimates indicate that the removal of the kinks did not significantly affect the shape of the distribution of daily wages above and below the kink.

Table B-2: RKD estimates at the 725SEK threshold: Inference

	(1) Unemployment Duration D	(2) Duration $D_{\mathbf{1}}$ (< 20 weeks)	(3) Duration $D_{\mathbf{2}}$ ($\geq$ 20 weeks)
	I. 1999-2000: Kink in b_1 and b_2		
Linear - $\delta_{\mathbf{k}}$	-.0569	-.0246	-.0299
Robust s.e.	(.0047)	(.0013)	(.0036)
Bootstrapped s.e.	(.0050)	(.0012)	(.0039)
95% CI - permutation test	[-.0595 ; -.0566]	[-.0319 ; -.0189]	[-.0402 ; -.019]
	II. 2001: Kink in b_2 only		
Linear - $\delta_{\mathbf{k}}$	-.0255	-.0115	-.0105
Robust s.e.	(.005)	(.0021)	(.0028)
Bootstrapped s.e.	(.0049)	(.0020)	(.0030)
95% CI - permutation test	[-.0325 ; -.0190]	[-.0127 ; -.0103]	[-.0115 ; -.0091]
	III. 2002-. : Placebo		
Linear -$\delta_{\mathbf{k}}$	.0045	-.0016	.006
Bootstrapped s.e.	(.0048)	(.0011)	(.0041)
Robust s.e.	(.0055)	(.0011)	(.0049)
95% CI - permutation test	[.0017 ; .0075]	[-.0021 ; -.0011]	[.0053 ; .0071]

Notes: The table reports estimates of the change in slope $\delta_{\mathbf{k}}$, at the 725SEK threshold, in the relationship between daily wage and the total duration of unemployment D (col. (1)), the time $D_{\mathbf{1}}$ spent on the first part of the Swedish UI profile (col. (2)) and the time $D_{\mathbf{2}}$ spent on the second part of the Swedish UI profile (col. (3)). $D_{\mathbf{1}} = \sum_{t<20wks} S_t$ corresponds to duration censored at 20 weeks of unemployment. $D_2 = \sum_{t\geq 20wks} S_t$ corresponds to unconditional duration spent unemployed after 20 weeks of unemployment (i.e., not conditional on having survived up to 20 weeks). Estimates are obtained from linear regressions of the form of equation (16) with a bandwidth size $h =$ 90SEK. These estimates are reported for three periods of interest. Panel I reports estimates for spells starting before July 2001. Panel II reports estimates for spells starting after July 2001 and before July 2002. Panel III reports estimates for spells starting after July 2002. Unemployment duration is capped at two years. We report for each estimate $\delta_{\mathbf{k}}$ the White robust standard errors, the bootstrapped standard errors using 50 replications, as well as 95% confidence intervals using the permutation test method of Ganong and Jaeger [2014].

Table B-3: RKD estimates at the 725SEK threshold: Sensitivity to polynomial order

	(1) Unemployment Duration D	(2) Duration D_1 (< 20 weeks)	(3) Duration D_2 ($\geq$ 20 weeks)
	I. 1999-2000: Kink in b_1 and b_2		
Linear - δ_k	-.0569	-.0246	-.0299
	(.0047)	(.0013)	(.0036)
RMSE	28.285	7.049	23.972
AIC	1785650.8	1264546	1723601.1
Quadratic - δ_k	-.0474	-.0344	-.0183
	(.0185)	(.0049)	(.0143)
RMSE	28.285	7.048	23.971
AIC	1785650.5	1264518.9	1723588.4
Cubic - δ_k	-.0527	-.0291	-.0221
	(.0455)	(.0122)	(.0351)
MSE	28.284	7.046	23.971
AIC	1785644.8	1264394.7	1723590
	II. 2001: Kink in b_2 only		
Linear - δ_k	-.0255	-.0115	-.0105
	(.0050)	(.0021)	(.0028)
RMSE	27.612	6.863	23.512
AIC	620999.2	438509.8	599929.6
Quadratic -δ_k	-.0579	-.0299	-.0151
	(.0196)	(.0098)	(.011)
MSE	27.612	6.863	23.512
AIC	621001.3	438509.9	599932.5
Cubic - δ_k	-.0192	-.0268	.0068
	(.0485)	(.0201)	(.0274)
MSE	27.612	6.863	23.512
AIC	621003.5	438508.6	599934.5

Notes: The table reports estimates of the change in slope δ_k, at the 725SEK threshold, in the relationship between daily wage and the total duration of unemployment D (col. (1)), the time D_1 spent on the first part of the Swedish UI profile (col. (2)) and the time D_2 spent on the second part of the Swedish UI profile (col. (3)). Estimates are obtained from polynomial regressions of the form of equation (16) with a bandwidth size $h = 90$SEK. Estimates are reported for three different polynomial orders: the linear specification, the quadratic specification and the cubic specification. For each polynomial order, we report model fit diagnostics: the root mean squared error (RMSE) as well as the Aikake information criterion (AIC). Panel I reports estimates for spells starting before July 2001. Panel II reports estimates for spells starting after July 2001 and before July 2002. White robust standard errors are in parentheses.

B.5 Exploiting Other Kinks to Assess the Robustness of $MH_1 \geq MH_2$

The Swedish system offers during the period 1999 to 2007 various sources of variations that can be used to identify $MH_{\mathbf{1}}$ and$MH_{\mathbf{2}}$. There are three simple reasons why the baseline estimates focus only on the kink at 725 SEK. First, for expositional convenience: since all statistics of interest can be identified using the same RK design, this made presenting the source of variation and the results particularly convenient. Second, for internal validity: comparing estimates at the same kink ensures that we are comparing behavioral responses for comparable individuals over time. And finally, for precision: there is more density around the 725SEK kink, which enables a higher degree of statistical precision for the RK estimates

In this section, we investigate the robustness of our results to the use of other sources of variations. Indeed, one may be worried that comparing individuals at the same kink over time may be problematic if behavioral elasticities are prone to varying over time (due to business cycle variations for instance). In this sense, there is a trade-off between comparing individuals at the same kink over time and comparing different individuals but in the same time period. In the first case, one may worry that time affects behavioral elasticities, while in the second case, one may worry that individuals at different kinks are different as they have different pre-unemployment incomes to start with.

In what follows, we review the different kinks and the sources of identification they provide, present our strategy and estimates for each kink, and summarize the conclusions that we can draw from this evidence on the relative magnitude of $MH_{\mathbf{1}}$ vs $MH_{\mathbf{2}}$. The bottom-line is that our estimates are very robust to using these alternative sources of identification and in particular the larger magnitude of $MH_{\mathbf{1}}$ relative to $MH_{\mathbf{2}}$ is a very robust finding, irrespective of the source of identification used.

B.5.1 850 SEK kink

In July 2001, the cap in $b_{\mathbf{1}}$, the UI benefits received during the first 20 weeks of unemployment, was increased to 680SEK, which created a kink in the relationship between $b_{\mathbf{1}}$ and the daily wage at a wage level of 850SEK. This gives us the possibility to identify in the RK design the effect of $b_{\mathbf{1}}$ on duration outcomes. In July 2002, the cap in $b_{\mathbf{2}}$, the UI benefits received after the first 20 weeks of unemployment, was increased to 680SEK, which created a kink in the relationship between $b_{\mathbf{2}}$ and the daily wage at a wage level of 850SEK.

Figure B-11 reports the evolution of the RKD estimates of the change in slope in the relationship between unemployment duration and daily wage at the 850 SEK kink, by year of entry into unemployment. Spells starting before 2001 are subject to a linear schedule with no kink in either $b_{\mathbf{1}}$ nor $b_{\mathbf{2}}$ and represent the placebo. Spells starting in 2001 are subject to a kink in $b_{\mathbf{1}}$ only. Spells starting in 2002 and after are subject to a kink in $b_{\mathbf{2}}$ only. The graph provides clear evidence of a break in the relative slopes of the relationship between duration and wage on both sides of the kink after the introduction of the kink in $b_{\mathbf{1}}$ in 2001. It also provides evidence of a slight decrease in the change in slope as the kink in $b_{\mathbf{1}}$ is replaced by a kink in $b_{\mathbf{2}}$ after 2002.

Based on this evidence, we implement a DD-RKD in order to get estimates of the elasticities of duration with respect to $b_{\mathbf{1}}$ and with respect to $b_{\mathbf{2}}$ at the 850SEK kink. Our DD-RKD specification is the following:

$$\begin{aligned} E[Y|w,P] = \quad & \beta_0 + \beta_1(w-\bar{w}) + \delta_0(w-\bar{w})\cdot \mathbb{1}[w \geq \bar{w}] \\ & + \mathbb{1}[P=1]\cdot\Big(\beta_2 + \beta_3(w-\bar{w}) + \delta_1(w-\bar{w})\cdot\mathbb{1}[w\geq\bar{w}]\Big) \\ & + \mathbb{1}[P=2]\cdot\Big(\beta_4 + \beta_5(w-\bar{w}) + \delta_2(w-\bar{w})\cdot\mathbb{1}[w\geq\bar{w}]\Big). \end{aligned} \tag{5}$$

where P denotes the time period in which the unemployment spell started. $P=0$ for spells starting before July 2001, $P=1$ for spells starting between July 2001 and July 2002 and $P=2$ for spells starting after July 2002.

From this specification, we compute the implied benefit elasticities of unemployment duration. The elasticity of duration with respect to $b_{\mathbf{1}}$, estimated at the 850SEK kink, is $\varepsilon_{D,b_{\mathbf{1}}} = \hat{\delta}_{\mathbf{1}} \cdot \frac{850}{D^{cap}}$, where $\hat{\delta}_{\mathbf{1}}$ is the estimated marginal slope change for spells starting in period $P=1$, and D^{cap} is the observed average duration at the kink for spells starting in period $P=1$. We find $\varepsilon_{D,b_{\mathbf{1}}} = 1.92$ (.31). The elasticity of duration with respect to b_2, estimated at the

850SEK kink, is $\varepsilon_{D,b_2} = \hat{\delta}_2 \cdot \frac{850}{D^{cap}}$, where $\hat{\delta}_2$ is the estimated marginal slope change for spells starting in period $P = 2$, and D^{cap} is the observed average duration at the kink for spells starting in period $P = 2$. We find $\varepsilon_{D,b_2} = 1.50 \ \ (.24)$.

B.5.2 912.5 SEK kink

In July 2002, the cap for b_1, benefits received during the first 20 weeks of unemployment, was increased to 730SEK, which created a kink in the schedule of benefits at a daily wage level of 912.5SEK. This gives us the possibility to identify the effect of b_1 on duration outcomes, using a RK design at the 912.5SEK kink.

As done previously at the 725SEK and 850SEK kink, Figure B-12 reports the evolution of the RKD estimates of the change in slope in the relationship between unemployment duration and daily wage at the 912.5SEK kink, by year of entry into unemployment. Spells starting before 2002 are subject to a linear schedule with no kink in either b_1 nor b_2 and represent the placebo. Spells starting in 2002 and after are subject to a kink in b_1 only. The graph provides clear evidence of a break in the relative slopes of the relationship between duration and wage on both sides of the kink after the introduction of the kink in b_1 in 2002. However, estimates are much less precise than for the other two kinks, as there is much less density around the 912.5SEK kink.

Based on this evidence, we also implement a DD-RKD in order to get estimates of the elasticities of duration with respect to b_1 at the 912.5SEK kink. Our DD-RKD specification is the following:

$$\begin{aligned} E[Y|w,P] = \quad & \beta_0 + \beta_1(w-\bar{w}) + \delta_0(w-\bar{w}) \cdot \mathbb{1}[w \geq \bar{w}] \\ & + \mathbb{1}[P=2] \cdot \Big(\beta_4 + \beta_5(w-\bar{w}) + \delta_2(w-\bar{w}) \cdot \mathbb{1}[w \geq \bar{w}]\Big). \end{aligned} \tag{6}$$

where P denotes the time period in which the unemployment spell started, with $P = 2$ for spells starting after July 2002.

From this specification, we compute the implied benefit elasticities of unemployment duration. The elasticity of duration with respect to b_1, estimated at the 912SEK kink is $\varepsilon_{D,b_1} = \hat{\delta}_2 \cdot \frac{912.5}{D^{cap}}$, where $\hat{\delta}_1$ is the estimated marginal slope change for spells starting in period $P = 2$, and D^{cap} is the observed average duration at the kink for spells starting in period $P = 2$. We find $\varepsilon_{D,b_1} = 2.15 \ \ (.70)$.

B.5.3 Combining estimates to identify the relative magnitude of MH_1 vs MH_2

Based on these sources of variations, we now have four different potential estimates of the relative moral hazard costs of b_1 vs b_2, summarized in Table B-4 below

The first strategy consists in comparing estimates at the same "kink" over time. This approach, as mentioned earlier, has the advantage of comparing similar individuals over time at the same level of income. One drawback may be that behavioral elasticities are time varying due to business cycle fluctuations for instance. Two kinks, the 725SEK kink and the 850SEK kink, allow us to implement this first strategy, giving us two different estimates of the moral hazard costs ratio, displayed in panel A and panel B of Table B-4. Results using estimates at the 725SEK kink represent our baseline strategy, displayed in panel A, and compare estimates from 1999-2000 versus 2001, taking advantage of the fact that at the 725SEK kink, we can identify $\varepsilon_{D,b}$ in 1999-2000 and ε_{D,b_2} in 2001. From this baseline strategy, we find a ratio $MH_1/MH_2 = 1.264$, as shown in panel A of Table B-4

Results using estimates at the 850SEK kink over time are displayed in panel B, and compare estimates from 2001 versus 2002-2007, taking advantage of the fact that at the 850SEK kink, we can identify ε_{D,b_1} in 2001 and ε_{D,b_2} in 2002-2007. From this strategy, we find a ratio $MH_1/MH_2 = 1.155$, as shown in panel B of Table B-4. Estimates of the moral hazard cost ratio from panel A and B are very similar. This confirm that the relative magnitude of the moral hazard costs is very robust across kinks, alleviating the concern that the moral hazard cost ratio is extremely sensitive to the location of the income distribution at which it is estimated.

The second strategy consists in comparing estimates at different "kinks" within the same time period. This second approach has the advantage of comparing individuals within the same time period, therefore controlling for the fact that behavioral elasticities may be time varying due to business cycle fluctuations for instance. A potential drawback of this approach is that individuals at different kinks may differ in their responsiveness to the policy.

Two time periods, 2001 and 2002-2007 allow us to implement this second strategy, giving us two additional estimates of the moral hazard costs ratio, displayed in panel C and panel D of Table B-4. For the 2001 period, we have estimates of $MH_{\mathbf{1}}$ from the 850SEK kink and estimates of $MH_{\mathbf{2}}$ from the 725SEK kink. Combining these estimates, we get a ratio $MH_{\mathbf{1}}/MH_{\mathbf{2}} = 2.695$, as shown in panel C of Table B-4. For the 2002-2007 period, we have estimates of $MH_{\mathbf{1}}$ from the 912.5SEK kink and estimates of $MH_{\mathbf{2}}$ from the 850SEK kink. Combining these estimates, we get a ratio $MH_{\mathbf{1}}/MH_{\mathbf{2}} = 1.343$, as shown in panel C of Table B-4. Again, both estimates of the MH cost ratio confirm that the moral hazard cost of $b_{\mathbf{1}}$ is larger than the moral hazard cost of $b_{\mathbf{2}}$ in our context.

Results from Table B-4 therefore all strongly suggest that our finding that $MH_{\mathbf{1}} > MH_{\mathbf{2}}$ is robust to the sources of variations used to estimate the effects of $b_{\mathbf{1}}$ and $b_{\mathbf{2}}$ on unemployment durations. We believe these results to be an important piece of additional evidence on the validity and robustness of our findings on the relative magnitude of moral hazard costs over the unemployment spell.

B.5.4 Inference on estimates of the relative magnitude of $MH_{\mathbf{1}}$ vs $MH_{\mathbf{2}}$

To provide inference on our estimates of the relative magnitude of $MH_{\mathbf{1}}$ vs $MH_{\mathbf{2}}$, we adopt a permutation-test approach. Since our estimation approach relies on a comparison of two kinks, either over time, or within period, we replicate this comparison for placebo kinks drawn at random in the distribution of daily wage in regions where the UI benefit schedule does not exhibit any kink.

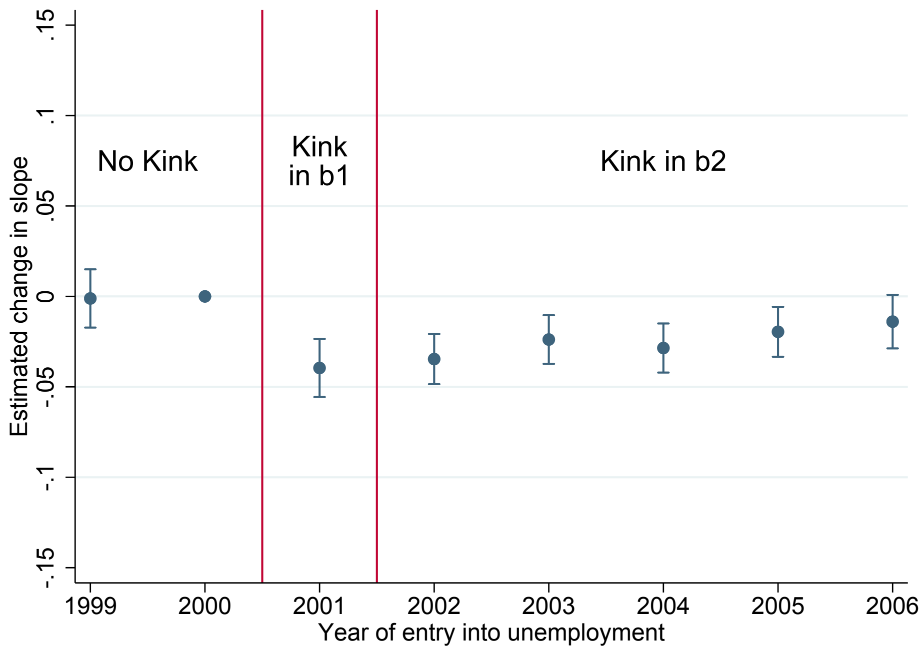
For panel A, we draw 250 kinks at random. For each kink, we estimate the change in slope in 1999-2000 and 2001 and from this obtain a distribution of 250 placebo estimates of the moral hazard $MH_{\mathbf{1}}$ and $MH_{\mathbf{2}}$. In Figure B-13 panel A, we draw the distribution of the placebo estimates of the difference $MH_{\mathbf{1}} - MH_{\mathbf{2}}$ obtained from this procedure. As can be seen, this distribution is centered around 0, and our true estimate obtained from the 725SEK kink lies far in the upper tail of this distribution of placebo estimates. This procedure gives us a p-value of 5.98% for our estimate of the relative magnitude of $MH_{\mathbf{1}} - MH_{\mathbf{2}}$. In other words, this procedure is providing compelling evidence that our estimates of the relative moral hazard costs are not picking up some random variation in the slope of the relationship between durations and daily wage over time.

Similarly, for panel B, we draw again 250 kinks at random. For each kink, we estimate the change in slope in 2001 vs 2002-2007 and from this obtain a distribution of 250 placebo estimates of the moral hazard $MH_{\mathbf{1}}$ and $MH_{\mathbf{2}}$. In Figure B-13 panel B., we draw the distribution of the placebo estimates of the difference $MH_{\mathbf{1}} - MH_{\mathbf{2}}$ obtained from this procedure. Again, this procedure confirms that our estimates of the relative moral hazard costs is not picking up some random variation in the slope of the relationship between durations and daily wage over time. From this distribution of placebo kinks, the p-value for our estimate in Table B-4 panel B is 0.00%.

Similarly, for panel C, we draw 250 kinks at random from which we can generate 31,000 ($= \frac{250^2}{2} - 250$) pairs of RKD estimates for the period July 2001 to July 2002. For each pair of kink, we estimate the changes in slope in 2001 and from this obtain a distribution of 31,000 placebo estimates of the moral hazard $MH_{\mathbf{1}}$ and $MH_{\mathbf{2}}$. In Figure B-13 panel C, we draw the distribution of the placebo estimates of the difference $MH_{\mathbf{1}} - MH_{\mathbf{2}}$ obtained from this procedure. Results again provide compelling evidence that our estimates of the relative moral hazard costs in panel C of Table B-4 are not picking up some random variation in the slope of the relationship between durations and daily wage across different kinks in the 2001 period. The p-value for our estimate in Table B-4 panel C that we obtain from this permutation procedure is 0.00%.

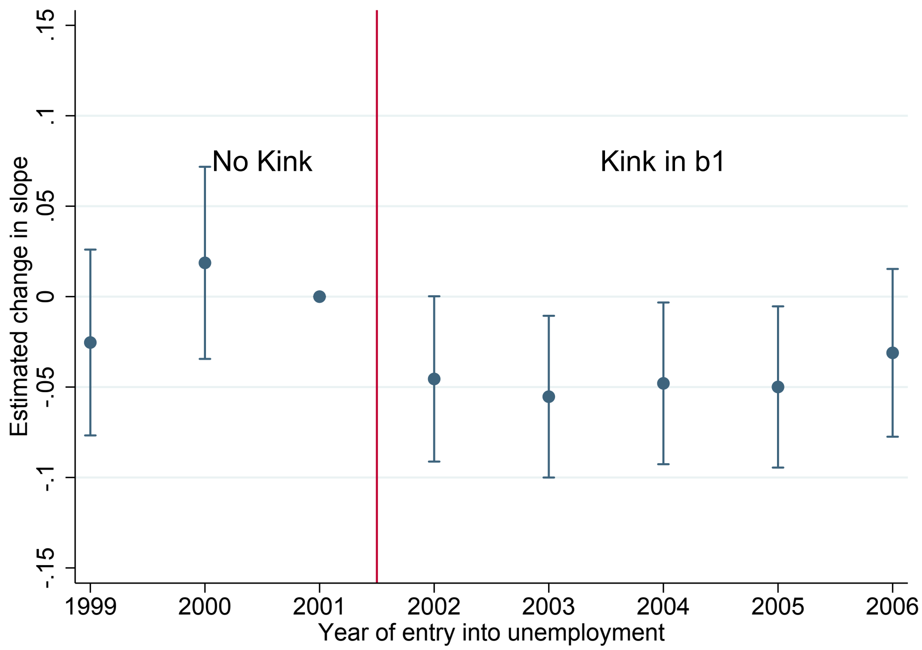
Finally, for panel D, we draw 250 kinks at random from which we can generate 31,000 ($= \frac{250^2}{2} - 250$) pairs of RKD estimates for the period 2002 to 2007. For each pair of kink, we estimate the changes in slope in 2002-2007 and from this obtain a distribution of 31,000 placebo estimates of the moral hazard $MH_{\mathbf{1}}$ and $MH_{\mathbf{2}}$. In Figure B-13 panel D., we draw the distribution of the placebo estimates of the difference $MH_{\mathbf{1}} - MH_{\mathbf{2}}$ obtained from this procedure. This gives us a p-value for our estimate in Table B-4 panel D of 2.63%.

Figure B-11: RKD estimates on unemployment duration D at the 850SEK kink by year of entry



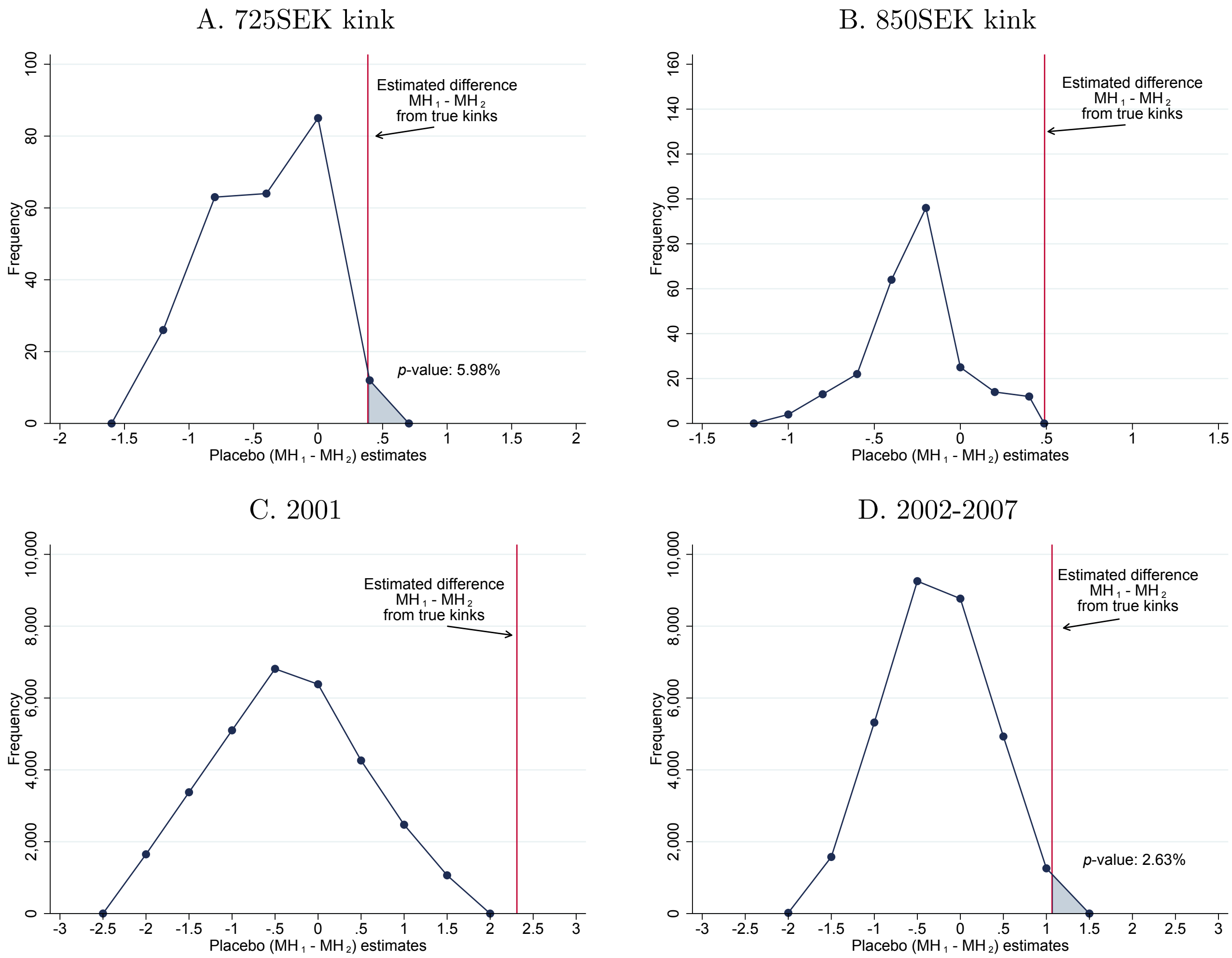
Notes: The figure reports the RKD estimates of the effect of UI benefits on total duration of unemployment by year of entry into unemployment, at the 850SEK kink. Entry into unemployment in Year N is defined as starting a spell between July 1st of Year N and July 1st of Year $N+1$. Spells starting before 2001 are therefore subject to a smooth linear schedule with no kink in either b_1 nor b_2 and represent the placebo. Spells starting in 2001 are subject to a kink in b_1 only. Spells starting in 2002 and after are subject to a kink in b_2 only. All estimates are from linear specifications using the changes in the UI schedule at the 850SEK kink with a 90SEK bandwidth, and are relative to year 2000, which is the baseline. 95% confidence interval around the point estimates, from robust standard errors, are displayed. The figure provides clear evidence that estimated responses in the RK design are indeed due to the policy changes, and not due to time trends in the distribution of durations around the kink.

Figure B-12: RKD estimates on unemployment duration D at the 912.5SEK kink by year of entry



Notes: The figure reports the RKD estimates of the effect of UI benefits on total duration of unemployment by year of entry into unemployment, at the 912.5SEK kink. Entry into unemployment in Year N is defined as starting a spell between of July 1st of Year N and July 1st of Year $N+1$. Spells starting before 2002 are subject to a smooth linear schedule with no kink in either b_1 nor b_2 and represent the placebo. Spells starting in 2002 and after are subject to a kink in b_1 only. All estimates are from linear specifications using the changes in the UI schedule at the 912.5SEK kink with a 60SEK bandwidth, and are relative to year 2001, which is the baseline. 95% confidence interval around the point estimates, from robust standard errors, are displayed.

Figure B-13: Permutation tests approach to inference on estimates of the relative magnitude of MH_1 vs MH_2



Notes: To provide inference on our estimates of the relative magnitude of MH_1 vs MH_2, we adopt a permutation-test approach. Since our estimation approach relies on a comparison of two kinks, either over time, or within period, we replicate this comparison for placebo kinks drawn at random in the distribution of daily wage in regions where the UI benefit schedule does not exhibit any kink. For each panel, we show the distribution of placebo estimates of the difference between moral hazard MH_1 and MH_2 from placebo kinks drawn at random using variation over time at the same kink (panel A, comparing 1999-2000 vs 2001 and panel B, comparing 2001 vs 2002-2007) or across kinks within the same period (panel C in 2001, and panel D in 2002-2007). The red vertical line in each panel displays the estimated value of the difference between moral hazard MH_1 and MH_2 from the true kinks (displayed in Table B-4).

Table B-4: Summary of Moral Hazard Cost Ratio Estimates

	I. Using Same Kink, Over Time	
	A. 725 SEK - Kink	B. 850 SEK - Kink
δ_{b_1}	-0.036	-0.060
	(0.007)	(0.009)
Variation used	*'99-'00 vs 2001 kink*	*2001 kink*
δ_{b_2}	-.0255	-0.047
	(.0049)	(0.007)
Variation used	*2001 kink*	*2002-2007 kink*
$\frac{MH_1}{MH_2}$	**1.264**	**1.155**
p-value	5.98%	0.00%
	II. Using Different Kinks Within Same Time Period	
	C. 2001	D. 2002-2007
δ_{b_1}	-0.060	-0.062
	(0.009)	(0.02)
Variation used	*850SEK kink*	*912.5SEK kink*
δ_{b_2}	-0.023	-0.047
	(0.008)	(0.007)
Variation used	*725SEK kink*	*850SEK kink*
$\frac{MH_1}{MH_2}$	**2.695**	**1.343**
p-value	0.00%	2.63%

Notes: The Table reports all the different estimates of the moral hazard cost ratios that can be drawn from the systematic exploitation of the kinks in the UI benefit schedule in Sweden over the period 1999 to 2007. Panel I reports estimates from the strategy consisting in comparing estimates at the same "kink" over time. This approach has the advantage of comparing similar individuals over time at the same level of income. Two kinks, the 725SEK kink and the 850SEK kink, allow us to implement this strategy, giving us two different estimates of the moral hazard costs ratio, displayed in panel A and panel B. Panel II reports estimates from the strategy consisting in comparing estimates at different "kinks" in the daily wage distribution within the same time period. This second approach has the advantage of comparing individuals within the same time period, therefore controlling for the fact that behavioral elasticities might be time varying due to business cycle fluctuations for instance. Two time periods, 2001 and 2002-2007 allow us to implement this second strategy, giving us two additional estimates of the moral hazard costs ratio, displayed in panel C and panel D. For each estimate of the moral hazard cost ratio, we report the RKD estimate of the effect on duration D of the change in slope for variation in b_1 (δ_{b_1}) and the estimate of the change in slope for variation in b_2 (δ_{b_2}), along with their standard errors in parenthesis. Below each RKD estimate δ_{b_1} and δ_{b_2}, we display the source of variation used for identification. Below each estimate of the MH cost ratio, we display p-values from a permutation based test, using placebo kinks drawn at random in the distribution of daily wage in regions where the UI benefit schedule does not exhibit any kink. See text for details.

B.6 Comparison to Duration Response Estimates in the Literature

How do our duration response estimates compare to existing estimates of labor supply responses to UI benefits available in the literature? To answer this question, we first need to be very precise about the source of variation of benefits used for the estimates against which we want to benchmark our estimates. As our conceptual analysis makes very clear, elasticity of duration w.r.t benefits paid early during the spell or w.r.t benefits paid later during the spell are conceptually different, and will likely be very different empirically, especially in the presence of non-stationary forces.

The elasticities $\varepsilon_{D_{\mathbf{k}},\bar{b}}$ that we report in Table 2 column (1) in the main text are unique, compared to estimates in the literature, because they measure duration responses to a unique form of benefit variation. To understand this, remember that the potential duration of benefits being infinite in the Swedish system during our period of analysis, individuals can collect these benefits indefinitely. This means that the elasticity $\varepsilon_{D_{\mathbf{k}},\bar{b}}$ that we report measures the response of duration to a change in benefits **forever**. We do not know of any other paper using similar source of variation. The reported elasticities with respect to this unique source of variation appear quite large at first glance, but it might not be that surprising, and it is hard to gauge these magnitudes using available benchmarks.

To appreciate the magnitude of our duration responses, it is therefore better to focus on elasticities that use variations in benefits that are similar to the ones used in the previous literature. Our elasticity of time spent in the first part of the profile $D_{\mathbf{1}}$ with respect to $b_{\mathbf{1}}$, the benefits that you receive in the first 20 weeks is, in this respect, probably the best candidate. It can for instance be compared to estimates of the elasticity of paid unemployment duration (unemployment duration up to exhaustion) with respect to a variation in UI "benefit level" in the US, given individuals receive these benefits only in the first 26 weeks in the US. Put it differently, the elasticity of paid unemployment duration with respect to a variation in UI "benefit level" in the US is conceptually equivalent to our estimate of $\varepsilon_{D_{\mathbf{1}},b_{\mathbf{1}}}$, with the slight difference that $b_{\mathbf{1}}$ is given for 20 weeks in Sweden instead of 26 weeks in the US, and D_1 is the duration up to 20 weeks while the duration of paid unemployment is the duration up to 26 weeks in the US.[4] Our estimate $\varepsilon_{D_{\mathbf{1}},b_{\mathbf{1}}}$ from the 725 SEK kink is .71 (.10). This estimate is in line with available estimates of the elasticity of paid unemployment duration with respect to a variation in UI "benefit level" in the US. The classic study is for instance Meyer [1990], who found an elasticity of .56.[5] Landais [2015] indeed finds a slightly smaller elasticity (around .4). Kroft and Notowidigdo [2016] find an elasticity of .63 (.33) at the average unemployment rate in the US (p.20, based on estimates in Table 2 column 1) and conclude that such estimate "is broadly similar to the previous literature (Moffitt 1985, Meyer 1990, Chetty 2008)".

This evidence suggests that when benchmarked against conceptually similar elasticities, our duration responses prove quite similar to moral hazard estimates available in the literature.

[4] Of course, even if they are conceptually similar, the two elasticities do not need to be the same. One obvious reason they may differ is that these elasticities are potentially endogenous to $b_{\mathbf{2}}$, and $b_{\mathbf{2}} = 0$ in the US, while $b_{\mathbf{2}} = b_{\mathbf{1}}$ in Sweden.

[5] See Meyer [1990], Table VI, column (7). Coefficient estimates for log(b) in the proportional hazard models can be interpreted as the elasticity of the hazard rate s with respect to the benefit level. However, under the assumption that the hazard rate is somewhat constant, these elasticities can be interpreted as elasticities of unemployment duration, since $D \approx 1/s$ so that $\varepsilon_D \approx -\varepsilon_s$

FOR ONLINE PUBLICATION - Appendix C: Additional results and robustness of consumption analysis

This Appendix presents the construction of our residual measure of consumption expenditures based on registry data. It also presents additional results on consumption patterns over the unemployment spell and dynamic selection. It finally compares our results to results obtained using household consumption surveys (HUT), for which we can also explore patterns of consumption over the spell across different categories of expenditures.

C.1 Registry-based measures of consumption

We start by describing the construction of a registry-based consumption expenditure measure and we explain how it can be used it to present complementary analysis of how consumption evolves during the unemployment spell.

The registry-based measure of consumption is based on exhaustive administrative information on income, transfers and wealth in Sweden accounting for all income sources and changes in assets. The measure offers the advantage of being computable for the universe of unemployed households from 1999 to 2007. The measure captures annual expenditures between December of each year.

We start from the accounting identity that expenditures in year n are the sum of all income and transfers received in period n, minus the change in assets between year $n-1$ and year n,

$$expenditures_n = income_n - \Delta assets_n.$$

As a result of the comprehensiveness of the longitudinal administrative dataset that we assembled including all earnings, income, taxes, transfers and wealth, we have precise third-party reported information on all the components needed to construct such residual measure of yearly expenditures for the universe of Swedish individuals and households for years 1999 to 2007.

Our approach builds on previous attempts to measure consumption from registry-data (e.g. Browning and Leth-Petersen [2003]) and in particular on Koijen et al. [2014] who constructed a similar measure in Sweden for years 2003 to 2007 using a smaller subset of individuals, and confirmed its consistency with HUT data. Our approach extends these previous attempts and is closely related to Eika et al. [2017] in exploiting additional information on asset portfolio choices and returns to reduce measurement error and excess dispersion of consumption measures based solely on first-differencing asset stocks.

For interested readers, Kolsrud et al. [2017], (in preparation for a special issue of the *Journal of Public Economics* dedicated to the use of new data sources for consumption analysis), offers all the details of the construction of our residual measure of consumption. In particular, it provides information on all the registers, variables and programs needed to construct a reliable registry-based measure of consumption from the Swedish administrative data, and explains how to use flow transaction information on assets (real estate assets and financial assets) rather than first-differences in asset stocks over time to reduce measurement error.

In practice, we compute consumption in year n as:

$$C_n = y_n + T_n + \tilde{C}_n^b + \tilde{C}_n^d + \tilde{C}_n^v + \tilde{C}_n^h,$$

where:

- y_n represents all earnings and is computed from the tax registers, which contain third-party reported earnings for all employment contracts, including all fringe benefits and severance payments.[6]
- T_n accounts for all income taxes and transfers, including unemployment insurance, disability insurance, sick pay, housing and parental benefits, etc.

[6] Note that self-employed, for whom earnings are in large part self-reported, are excluded from the analysis as they are part of a different UI system.

- $\tilde{C}_n^b = y_n^b - \Delta b_n$ equals consumption out of bank holdings. It is equal to interests earned on these bank holdings during year n, y_n^b, minus the change in the value of bank holdings between year $n-1$ and year n, $\Delta b_n = b_n - b_{n-1}$.
- $\tilde{C}_n^d = -y_n^d + \Delta d_n$ is consumption out of debt, which includes student loans, credit card debt, mortgages, etc., and is third-party reported by financial institutions to the tax authority. It is equal to the change in the stock of debt Δd_n, minus all interests paid on the existing stock of debt y_n^d.
- $\tilde{C}_n^v = y_n^v - \Delta v_n$ is consumption out of financial assets (other than liquid holdings in bank accounts). It is equal to all income from financial assets y_n^v minus the change in the value of the portfolio of financial assets Δv_n. The return on financial assets y_n^v includes interests, dividends and any price change $\Delta p_n^v \times q_n^v$. Such price change would be exactly offset by a change in the value of assets, included in Δv_n, unless the return is realized by selling the asset. In practice, to separate the contribution of price changes from contribution of asset rebalancing, we use detailed data collected by Statistics Sweden from banks and financial institutions on all financial securities held by individuals, which contain information on quantities, ISIN numbers, and transaction values and dates.
- $\tilde{C}_n^h = y_n^h - \Delta h_n$ constitutes consumption out of real estate wealth. It is equal to all income derived from holding real estate assets y_n^h minus the change in the value of real estates Δh_n. Detailed information on the stock of real estate wealth, estimated at market value as of December 31 of each year, is available from the tax authority. The return to holding real estate y_n^h includes rents, but also imputed rents for homeowners, as well as price changes in the value of real estates. Such price change would again be exactly offset by a change in the value of real estate assets, included in Δh_n, unless the return is realized by selling the asset. To separate the contribution of price changes from contribution of real estate buying/selling, we use the housing transaction register which collects information on all real estate transactions operated in Sweden.

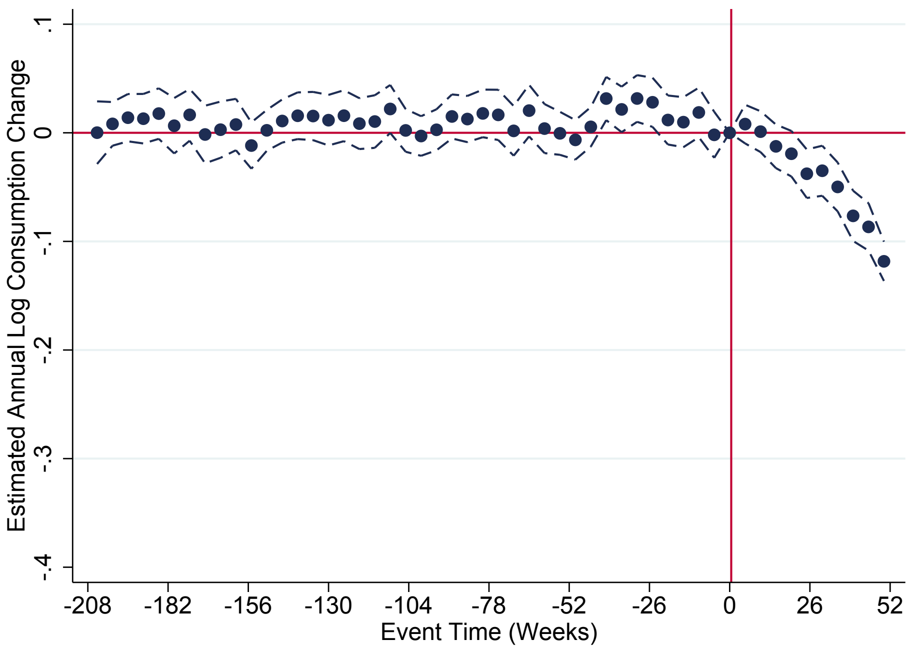
All income, transfers and asset positions are reported to the tax administration (and observed in the data) at the individual level.[7] We have aggregated consumption measures at the household level using household identifiers constructed by the Swedish National Statistical Office (Statistics Sweden).

C.2 Consumption patterns prior to the onset of a spell

To recover the drops in the first and second part of the profile from drops in annual consumption expenditures, our baseline implementation in section 4.1 makes the assumption that $c_j = \bar{c}_0, \forall j < 0$, or in other words, that the consumption profile is flat prior to the start of an unemployment spell. To investigate how consumption profiles evolve prior to the unemployment spell we report in Figure C-1 the annual log household consumption changes as a function of time t since the onset of a spell, $\mathscr{C}_{i,t} - \mathscr{C}_{i,t-52}$ in 5 weeks bins, relative to the last 5 weeks prior to the onset of a spell. We go as far back as 4 years prior to the onset of a spell, to detect longer anticipation effects. The Figure shows two interesting patterns. First, the consumption patterns of households are extremely stable in the 4 years prior to displacement, suggesting no long term anticipation behaviors. Second, there does not seem to be any sharp consumption changes immediately preceding displacement, suggesting no significant short term anticipation behaviors. Overall, this evidence strongly supports the assumption that flow consumption profiles are flat prior to the onset of an unemployment spell.

[7] It is possible that our measure misses some financial wealth held in foreign banks to the extent that these banks do not comply with the requirement to transmit information on the financial wealth of their Swedish customers to the Swedish tax authority. Yet, the fraction of foreign-held assets is low (about 3% of household assets according to the Savings Barometer of Statistics Sweden), and likely to be held by much wealthier households than our sample of unemployed.

Figure C-1: Change in log annual consumption as a function of time since onset of an unemployment spell: pre-unemployment trends



Notes: The Figure investigates the validity of the assumption of a flat profile of consumption pre-unemployment. We report the annual log household consumption changes as a function of time t since the onset of a spell, $\ln \mathscr{C}_{i,t} - \ln \mathscr{C}_{i,t-52}$ in 5 weeks bins, relative to the last 5 weeks prior to the onset of a spell. Estimates provide compelling evidence that consumption profiles are flat prior to the onset of an unemployment spell.

C.3 Additional Results on Consumption Smoothing Means & Dynamic Selection

Assets and liquidity constraints In this subsection, we discuss the evidence of the role of assets and liquidity constraints in consumption smoothing over the spell. First, we report in Appendix Figure C-2, and following the same methodology as in Figure 6, the average annual consumption drops by time spent unemployed as of December, breaking down the sample by the level of net wealth of the household at the start of the spell. The Figure shows that households with higher net wealth experience a lower drop in consumption conditional on unemployment duration. These households tend to have a smoother consumption profiles during unemployment, especially earlier on in the spell.

Second we provide direct evidence that individuals do use their liquid assets to smooth consumption over the spell. In Appendix Figure C-3 panel A, we estimate average change in liquid bank holdings compared to pre-unemployment by time t spent unemployed as of December (when bank holding stock is observed in the registry data). Estimates of the change in bank holdings $\mathscr{B}$ are scaled by the average annual consumption in the last year prior to unemployment, so that all changes in bank holdings are expressed relative to pre-unemployment household consumption levels. The Figure provides evidence that households use their liquid assets to smooth consumption over the unemployment spell, by depleting their bank accounts or reducing their savings. This source of consumption smoothing remains small in magnitude, as even after more than a year of unemployment, the change in bank holdings represent around 2% of pre-unemployment consumption levels of the household.

In panel B of Appendix Figure C-3 we report similar estimates for all other assets. The Figure shows that the change in total net assets other then bank holdings represent $\approx$2.5% of pre unemployment consumption after a year in unemployment.

Finally, evidence from registry data also indicates that debt does not offer much help in smoothing consumption over the unemployment spell, hinting at the presence of liquidity constraints. In Appendix Figure C-4, we provide evidence of a reduction in the use of non mortgage-related credit over the unemployment spell among households with no real estate and no mortgage debt. This Figure shows average change in debt compared to pre-unemployment by time t spent unemployed as of December (when debt level is observed in the registry data). The sample is restricted to individuals with no real estate wealth throughout the sample period. Because we cannot precisely separate mortgage debt from other types of credit in the data, this sample restriction is a direct way to identify how non-mortgage related debt evolves over the unemployment spell. The Figure provides evidence that households reduce their debt level rather than increase it as they become unemployed. Instead of contributing positively to consumption smoothing, debt contributes negatively to consumption over the unemployment spell. This suggests that as the duration of the spell increases, access to credit becomes harder and consumption out of debt falls significantly.

Overall, this analysis confirms that households may have means to smooth consumption over short spells. But as the duration of the spell increases, these means get quickly exhausted. Households in the second part of the profile seem therefore closer to being hand-to-mouth.

Spousal Labor Supply Within a household, the labor supply of other members of the household may help reduce the drop in household consumption over the spell. In Figure C-5 we investigate how labor supply of other members of the household affects the drop in household consumption over the spell. Following the same methodology as in Figure 6, we report the average change in total gross earnings and in total disposable income of all other members of the household as a function of time spent unemployed, scaled by the annual household consumption level prior to unemployment. Individual disposable income include all individual taxes and transfers, as well as individual capital gains and losses.

Results show that within-household changes in the earnings and disposable income of all other members of the household are small and almost always not significantly different from zero throughout the unemployment spell. This suggests that in our context, the labor supply of other members of the household does not play a significant role in smoothing household consumption even for long-term unemployed. This may be because the added-worker effect is not playing a significant role in increasing household consumption in response to an unemployment shock, or because its effect is mitigated by correlated earnings shocks across household members.

Dynamic Selection In practice, agents are heterogeneous and selection into longer unemployment spells may affect consumption responses over the spell and the gradient of consumption smoothing gains. To assess the potential magnitude of dynamic selection on consumption and risk preferences, we investigate in Table C-1 how various observable characteristics that have been shown to correlate with consumption and risk preferences are distributed across short term and long term unemployed.

To do so, we restrict the sample to all individuals about to become unemployed in the next quarter and estimate a linear probability model where the outcome is an indicator variable for experiencing a future spell longer than 20 weeks. The default age category is 18 to 30 years old. Income refers to individual disposable income and results are relative to the first quintile. Net wealth results are relative to individuals with zero or negative net wealth at the start of the spell. We also investigate the effect of two portfolio characteristics, that, conditional on net wealth, are traditionally correlated with risk preferences. First, we look at the fraction of total wealth invested in stocks, and results are relative to the first two quartile of this distribution (50% of the sample have zero stocks prior to becoming unemployed). Second, we look at leverage defined as total debt divided by gross assets, and results are relative to the first quartile of leverage.

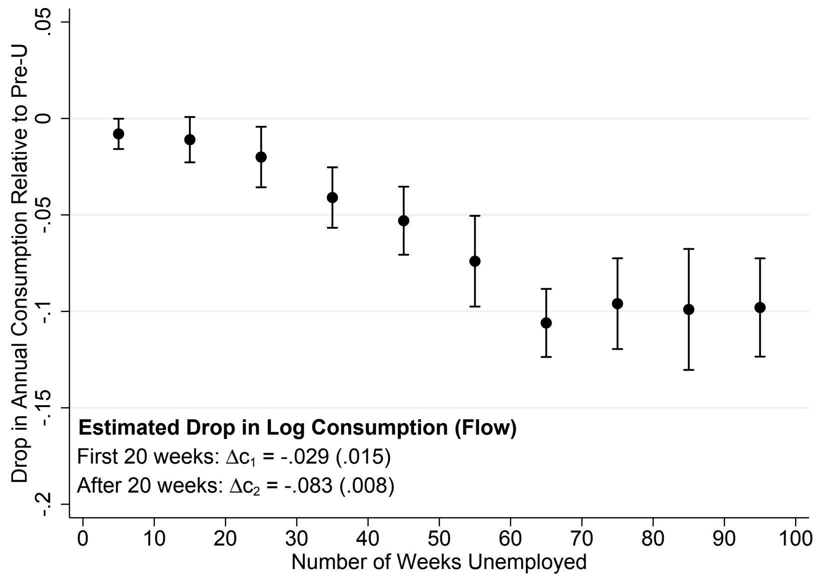
The patterns of selection are generally small in magnitude and often ambiguous in sign. Income levels and net wealth levels have quantitatively small and non-monotonic effects on the probability to select into longer spells. Compared to households with no or negative net wealth, households with some small wealth (<500kSEK) have a slightly lower probability to be long term unemployed. But individuals with high net wealth have a slightly higher probability to select into long spells. This result suggests that in our context, individuals with better means to smooth consumption do not unambiguously select into longer spells, which corroborates the evidence of no clear patterns of selection on consumption profiles displayed in subsection 4.1. Also portfolio characteristics (i.e., the fraction of portfolio wealth invested in stocks, and leverage defined as total debt divided by gross assets), which have been shown to be correlated with risk preferences, have small and non-monotonic impacts on the probability to experience a long unemployment spell. In contrast, the probability of experiencing long unemployment spells ($\mathbb{1}[L > 20 \text{ wks}]$) is significantly and monotonically correlated with age. However, existing evidence from the literature suggests a U-shape relationship between age and risk aversion (Cohen and Einav [2007]), so the dynamic selection on age has an ambiguous effect on the evolution of risk preferences over the spell.

Overall, the observed dynamic selection patterns on consumption and risk preferences do not suggest that relative consumption smoothing gains would be significantly different than our estimates based on the average drops in consumption.[8]

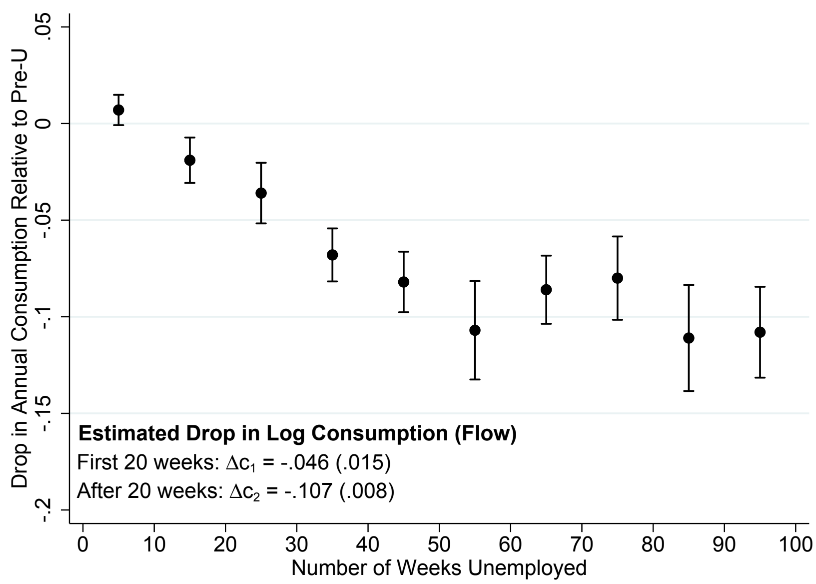
[8]Yet, other methods exploiting comparative statics of effort choices as in Chetty [2008a], or Landais [2015], could also be developed to evaluate the evolution of consumption smoothing gains over the spell. These methods could circumvent the issue of having to make assumptions regarding dynamic selection on risk preferences.

Figure C-2: Estimated Drop in Annual Consumption Relative to Pre-Unemployment as a Function of Time Spent Unemployed: Heterogeneity By Level of Wealth

A. Above Median Wealth Level Pre-Unemployment

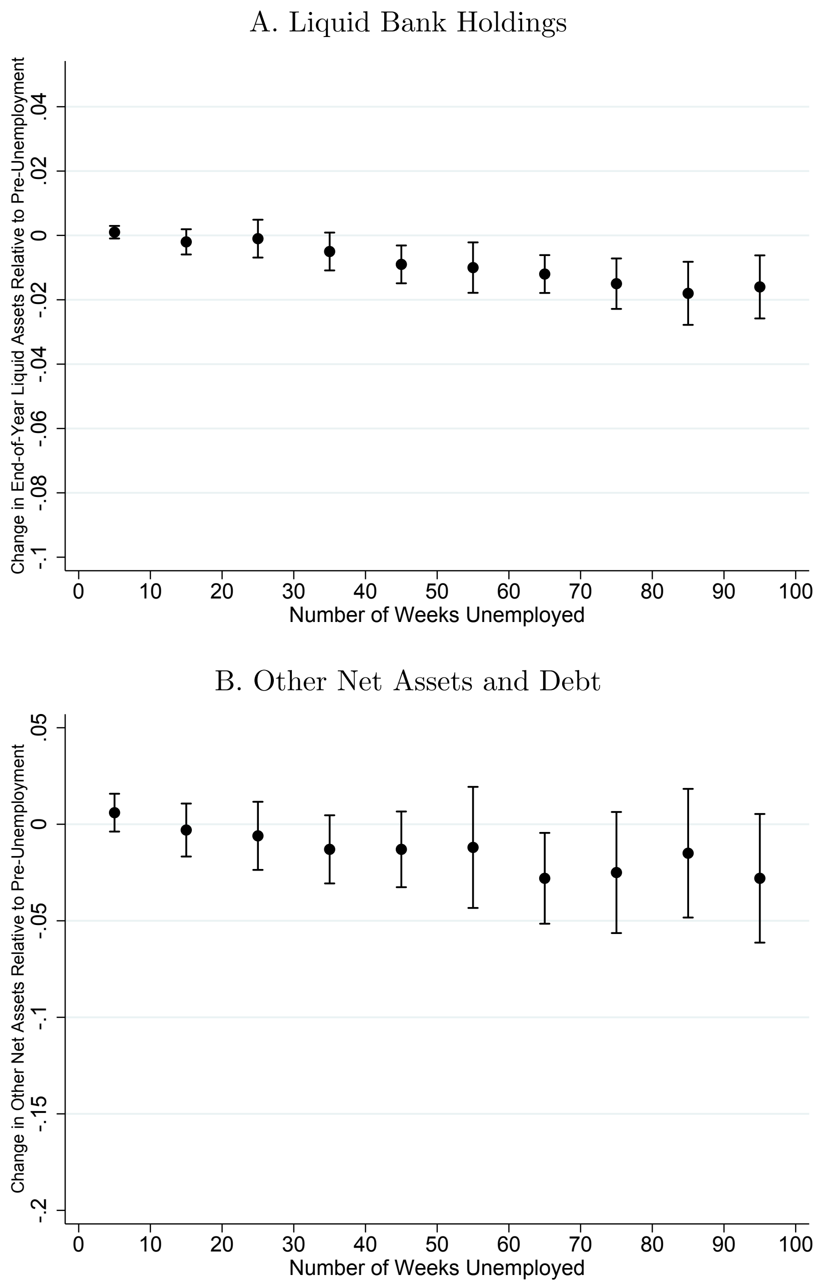


B. Below Median Wealth Level Pre-Unemployment



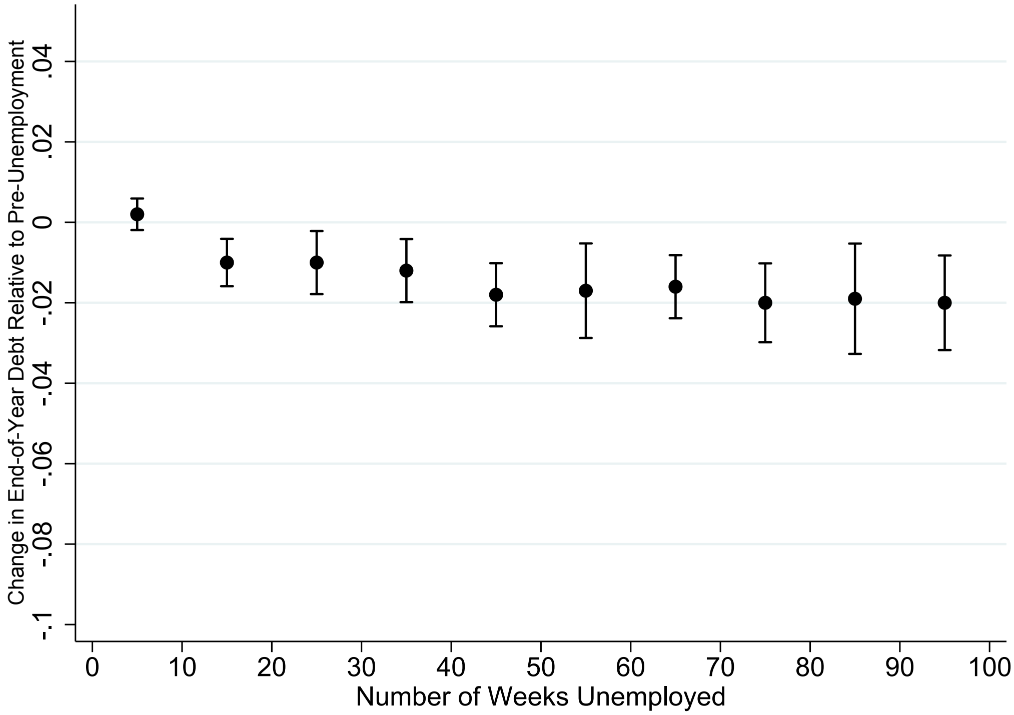
Notes: This Figure shows average annual consumption drops compared to pre-unemployment $\Delta\mathscr{C}_t$ by time t spent unemployed as of December (when annual consumption is observed in the registry data), and compares consumption profiles for households according to their level of net wealth prior to becoming unemployed. In both panels, we report $\hat{\beta}_t/\mathscr{C}_0 = \widehat{\Delta\mathscr{C}_t}/\mathscr{C}_0$, i.e. estimates from equation (20) scaled by the average annual consumption in the last year prior to unemployment, so that all consumption drops are expressed relative to pre-unemployment levels. The Figure also reports non-parametric estimates of the average drops in consumption in each parts of the benefit profile $\widehat{\Delta C_1}$ and $\widehat{\Delta C_2}$ following the methodology explained in subsection 4.1. Standard errors are computed using the Delta-method. Panel A reports estimates for households whose net wealth level in the last year prior to becoming unemployed is above the median wealth level in the sample. Panel B reports estimates for households whose net wealth level in the last year prior to becoming unemployed is below the median wealth level in the sample. The Figure provides evidence that households with higher net wealth experience smaller consumption drops conditional on unemployment duration.

Figure C-3: Estimated Change in Liquid Bank Holdings and Other Net Assets Relative to Pre-Unemployment as a Function of Time Spent Unemployed



Notes: This Figure shows average change in household liquid bank holdings $\mathscr{B}$ (panel A.) and other net assets $\mathscr{A}$ (panel B.) compared to pre-unemployment by time t spent unemployed as of December (when bank holding stock and other assets are observed in the registry data). We follow the same specification as for consumption drops (from equation (20)). In panel A, we report $\hat{\beta}_t/\mathscr{C}_0 = \widehat{\Delta\mathscr{B}_t}/\mathscr{C}_0$, i.e. estimates of the change in bank holdings $\mathscr{B}$ scaled by the average annual consumption in the last year prior to unemployment, so that all changes in bank holdings are expressed relative to pre-unemployment household consumption levels. In panel B we report $\hat{\beta}_t/\mathscr{C}_0 = \widehat{\Delta\mathscr{A}_t}/\mathscr{C}_0$. The Figure provides evidence that households use their assets to smooth consumption over the unemployment spell. This source of consumption smoothing remains small in magnitude. After more than a year of unemployment, the change in bank holdings represent 2% of pre-unemployment consumption levels of the household. Other net assets provide an additional 2.5% consumption smoothing after a year of unemployment.

Figure C-4: Estimated Change in Non-Mortgage Debt Relative to Pre-Unemployment as a Function of Time Spent Unemployed

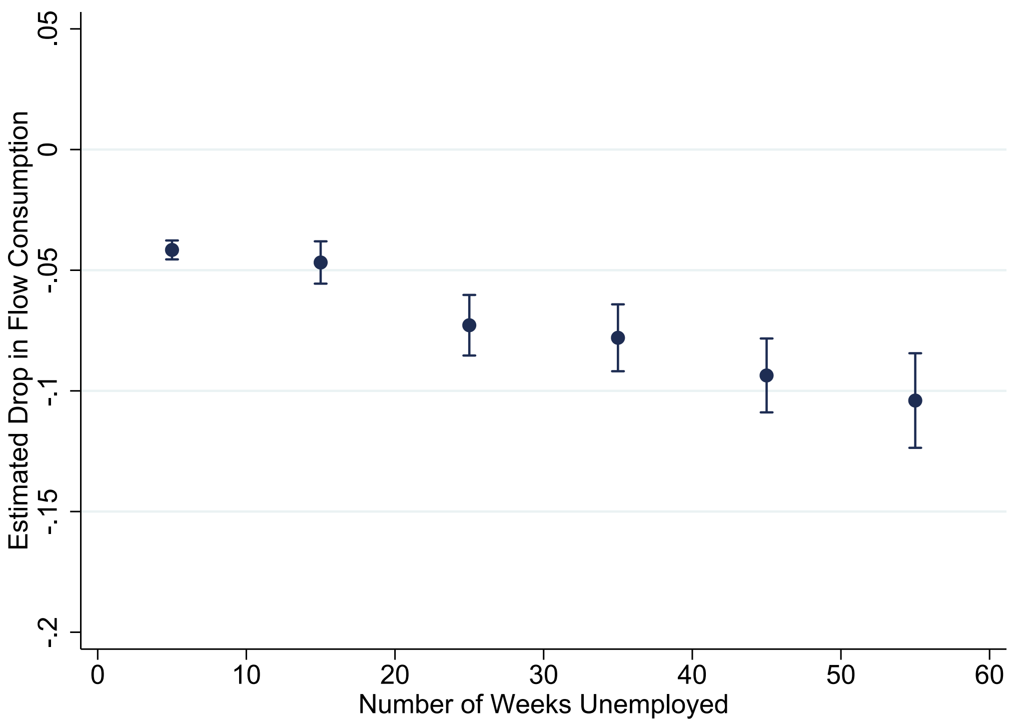


Notes: This Figure shows average change in debt compared to pre-unemployment by time t spent unemployed as of December (when debt level is observed in the registry data). The sample is restricted to individuals with no real estate wealth throughout the sample period. Because we cannot precisely separate mortgage debt from other types of credit in the data, this sample restriction is a direct way to identify how non-mortgage related debt evolves over the unemployment spell. We follow the same specification as for consumption drops (from equation (20)) and report $\hat{\beta}_t/\mathscr{C}_0 = \widehat{\Delta \mathscr{D}_t}/\mathscr{C}_0$, i.e. estimates of the change in debt $\mathscr{D}$ scaled by the average annual consumption in the last year prior to unemployment, so that all changes in debt are expressed relative to pre-unemployment household consumption levels. The Figure provides evidence that households reduce their debt level rather than increase it as they become unemployed. Instead of contributing positively to consumption smoothing, debt contributes negatively to consumption over the unemployment spell, which is suggestive of the presence of credit constraints among unemployed individuals.

Figure C-5: Estimated Change in Earnings and Disposable Income of all other members of the household as a function of unemployment duration

Notes: This figure explores the role of other household members (i.e. all members of the household excluding the unemployed individual) in smoothing household consumption over the spell. We follow the same specification as for consumption drops (from equation (20)) and report $\hat{\beta}_t/\mathscr{C}_0 = \widehat{\Delta\mathscr{Y}_t}/\mathscr{C}_0$, i.e. estimates of the change in gross earnings or disposable income $\mathscr{Y}$ of all the other household members scaled by the average annual consumption in the last year prior to unemployment, so that all changes are expressed relative to pre-unemployment household consumption levels. Disposable income includes individual taxes and transfers, as well as capital gains/losses. Results show that within-household changes in earnings and disposable income of all other members of the household are extremely small throughout the unemployment spell.

Figure C-6: Estimated drop in flow consumption over the first year of the unemployment spell



Notes: The Figure reports non-parametric estimates of the average drops in flow consumption at 10 week frequency following the methodology explained in subsection 4.1. Standard errors are computed using the Delta-method.

Table C-1: Pre-unemployment characteristics of individuals with spells longer than 20 weeks. Linear probability model estimates

	(1)	(2)	(3)	(4)	(5)
	Duration of future spell ≥ 20 weeks				
Age: 30 to 39	0.129***	0.118***	0.116***	0.119***	0.120***
	(0.00237)	(0.00250)	(0.00251)	(0.00305)	(0.00311)
Age: 40 to 49	0.164***	0.153***	0.153***	0.162***	0.163***
	(0.00277)	(0.00293)	(0.00295)	(0.00357)	(0.00363)
Age: 50+	0.272***	0.261***	0.265***	0.281***	0.282***
	(0.00288)	(0.00307)	(0.00319)	(0.00367)	(0.00371)
Married	0.0289***	0.0283***	0.0287***	0.0185***	0.0190***
	(0.00243)	(0.00243)	(0.00244)	(0.00280)	(0.00281)
Gender: Female	-0.00226	-0.00209	-0.00279	-0.0146***	-0.0135***
	(0.00192)	(0.00193)	(0.00193)	(0.00230)	(0.00230)
# of children	-0.0329***	-0.0315***	-0.0288***	-0.0318***	-0.0311***
	(0.00193)	(0.00193)	(0.00193)	(0.00231)	(0.00233)
2nd quintile of income		0.0412***	0.0436***	0.0321***	0.0319***
		(0.00319)	(0.00319)	(0.00409)	(0.00412)
3rd quintile of income		0.0842***	0.0885***	0.0850***	0.0849***
		(0.00319)	(0.00319)	(0.00403)	(0.00415)
4th quintile of income		0.0471***	0.0532***	0.0532***	0.0545***
		(0.00328)	(0.00329)	(0.00404)	(0.00421)
5th quintile of income		0.0453***	0.0518***	0.0589***	0.0635***
		(0.00341)	(0.00345)	(0.00411)	(0.00431)
0<Net wealth≤200k			-0.0503***	-0.0116***	-0.0122***
			(0.00234)	(0.00271)	(0.00315)
200k<Net wealth≤500k			-0.0466***	-0.0146***	-0.0114***
			(0.00324)	(0.00350)	(0.00425)
500k<Net wealth≤5M			-0.0186***	0.00576*	0.00774*
			(0.00300)	(0.00336)	(0.00418)
Net wealth>5M			0.0731***	0.0852***	0.0866***
			(0.0173)	(0.0172)	(0.0174)
Fraction of portfolio in stocks					
3rd quartile				-0.000542	
				(0.00787)	
4th quartile				0.0303***	
				(0.00259)	
Leverage: debt / assets					
2nd quartile					0.0153***
					(0.00390)
3rd quartile					-0.0120***
					(0.00322)
4th quartile					-0.00629*
					(0.00361)
R^2	0.0465	0.0490	0.0511	0.0624	0.0620
N	269931	269931	269931	190176	190176

Notes: The Table assesses the robustness of our welfare conclusions to dynamic selection on risk preferences over the unemployment spell. We investigate how various observable characteristics correlate with the probability to experience a long unemployment spell. To do so, we restrict the sample to all individuals about to become unemployed in the next quarter and estimate a linear probability model where the outcome is an indicator variable for experiencing a future spell longer than 20 weeks. The default age category is 18 to 30 years old. Income refers to individual disposable income and results are relative to the first quintile. Net wealth results are relative to individuals with zero or negative net wealth at the start of the spell. We also investigate the effect of two portfolio characteristics, that, conditional on net wealth, are traditionally correlated with risk preferences. First, we look at the fraction of total wealth invested in stocks, and results are relative to the first two quartile of this distribution (50% of the sample have zero stocks prior to becoming unemployed). Second, we look at leverage defined as total debt divided by gross assets, and results are relative to the first quartile of leverage.

C.4 Consumption profiles over the spell using household consumption surveys (HUT)

In order to analyze the evolution of consumption as a function of time spent unemployed, we have also merged the household consumption surveys (HUT) with the universe of administrative UI records. This enables us to reconstruct the full employment history of individuals whose household is surveyed in the HUT. We observe employment status of all individuals prior, during and after their HUT interview.

We restrict the sample to households where an individual is either unemployed at the time of the interview, or who will become unemployed some time in the next two years following the interview. This leaves us with a pseudo-panel of households for which we can correlate flow measures of consumption with time since (or until) the onset of the unemployment spell. Note that this sample is a pseudo-panel and not a panel *stricto sensu*, as households are surveyed only once in the HUT. However, because we observe the full unemployment history of individuals irrespective of the time they are surveyed, we can fully control for selection issues arising from differences between households who select into spells of different lengths.

In Table C-2, we provide summary statistics for unemployed individuals from the HUT sample. Unemployment duration patterns are almost identical in the HUT sample and in the RKD sample, with an average duration of unemployment of 26.6 weeks, and an average time spent in the first (resp. second) part of the profile of 12.9 weeks (resp. 13.8 weeks). Average replacement rates are also identical at 72%. Interestingly, unemployed individuals have similar demographics in the two samples. The distribution of age, gender, marital status and education levels are almost identical. Average earnings are also similar in the two samples, but the distribution of earnings is by construction much more spread out in the HUT sample than in the RKD sample, because the RKD sample is focused on unemployed with daily wages in the neighborhood of 725SEK. Finally, the two samples have very comparable distribution of assets and debts. In both samples, unemployed individuals have a household net wealth equal to 2.5 times their yearly earnings at the start of the spell. Overall, samples of unemployed individuals in the HUT and RKD analysis are well-balanced in terms of all observable characteristics and unemployment durations.

We start by providing graphical evidence for the evolution of average consumption as a function of time spent unemployed. We report in Figure C-7 the estimated coefficients β_k from the regression model:

$$c_{i,t} = \sum_{k=-3}^{+4} \beta_k \cdot \mathbb{1}[t \in [26 * k, 26 * \{k+1\})] + X_i'\gamma + \varepsilon_{i,t} \tag{7}$$

where $c_{i,t}$ is log household consumption in the HUT survey observed at the time of the HUT interview. t is event time in weeks since the start of the unemployment spell. We aggregate event time in bins of 6 months periods (26 weeks), so that $\mathbb{1}[t \in [26 * k, 26 * \{k+1\})]$ is an indicator for being observed in the k-th 6 months period since the start of an unemployment spell.[9] We include in this regression a set of controls X, which consists of year dummies, calendar month dummies and a set of dummies for family status. The estimated coefficients β_k plotted in Figure C-7 represent the average log household consumption for households with an individual observed in her $26 * k$-th to $26 * (k+1)$-th week of unemployment, relative to the average consumption level of households with an individual observed just prior to becoming unemployed.

The graph provides strong evidence corroborating our findings from the registry-based measure of consumption, that average household consumption drops significantly when unemployed. The average consumption of households where a member has been unemployed for more than a year is almost 15% lower than prior to the unemployment spell. Furthermore, Figure C-7 indicates that consumption declines significantly throughout the unemployment spell. The consumption drop is still limited early on in the spell, but average consumption decreases sharply as duration increases, and seems to be reaching a plateau after about a year in the unemployment spell. The last interesting finding is the lack of anticipation prior to the unemployment spell. Household consumption is flat in the year

[9] $k = 0$ is the baseline and represents consumption in the last 6 months prior to becoming unemployed. HUT surveys collect information on bi-weekly household consumption expenditures at the time of the interview. $c_{i,t}$ is then annualized by multiplying bi-weekly consumption by 26.

preceding the onset of the unemployment spell. This also confirms evidence from the registry data and suggests that unemployment shocks are relatively unanticipated, or that households have little propensity to change consumption prior to becoming unemployed to smooth out the upcoming earnings shock.

To recover the average consumption drops in the first and second part of the unemployment spell and to investigate the role of selection, we move in Table C-3 to a regression setting. In column (1), we start by running the simple regression:

$$c_{i,t} = \eta_1 \cdot \mathbb{1}[0 < t <= 20 \text{ wks}] + \eta_2 \cdot \mathbb{1}[t > 20 \text{ wks}] + X_i'\gamma + \varepsilon_{i,t}, \tag{8}$$

where $c_{i,t}$ is log consumption of the household and η_1 (resp. η_2) captures the effect of having a member of the household unemployed for less than 20 weeks (resp. more than 20 weeks) at the time of the interview, relative to households where one individual will become unemployed in the next 6 months. The baseline specification reported in column (1) controls for year dummies and calendar month dummies. Results suggest that relative to pre-unemployment levels, total average household consumption drops by 6% in the first 20 weeks of unemployment, and by 13% after 20 weeks of unemployment.

In column (2) of Table C-3, we add non-parametric controls for marital status, household size and age, and find similar results with an initial drop of 4% in the first 20 weeks and 13% after 20 weeks of unemployment. In columns (3) and (4), we analyze the role of selection in explaining these consumption patterns. If households with long unemployment spells are inherently different from households with short unemployment spells, the estimated drops in consumption after 20 weeks in column (1) and (2) may partly pick up these differences. To investigate such selection effects, we estimate in column (3) a model of the form

$$c_{i,t} = \eta_1 \cdot \mathbb{1}[0 < t <= 20 \text{ wks}] + \eta_2 \cdot \mathbb{1}[t > 20 \text{ wks}] + \alpha \cdot \mathbb{1}[L > 20 \text{ wks}] + X_i'\gamma + \varepsilon_{i,t}, \tag{9}$$

where $\mathbb{1}[L > 20 \text{ wks}]$ is a dummy variable for the completed length of the unemployment spell being longer than 20 weeks. In other words, because we observe in the UI records the total duration of each spell irrespective of the time when the household is surveyed in the HUT, we can control for consumption differences at any point in time between households with long versus short spells. Results indicate that between households who select into spells longer than 20 weeks versus households who select into spells shorter than 20 weeks, differences in log consumption levels are small and not significant.

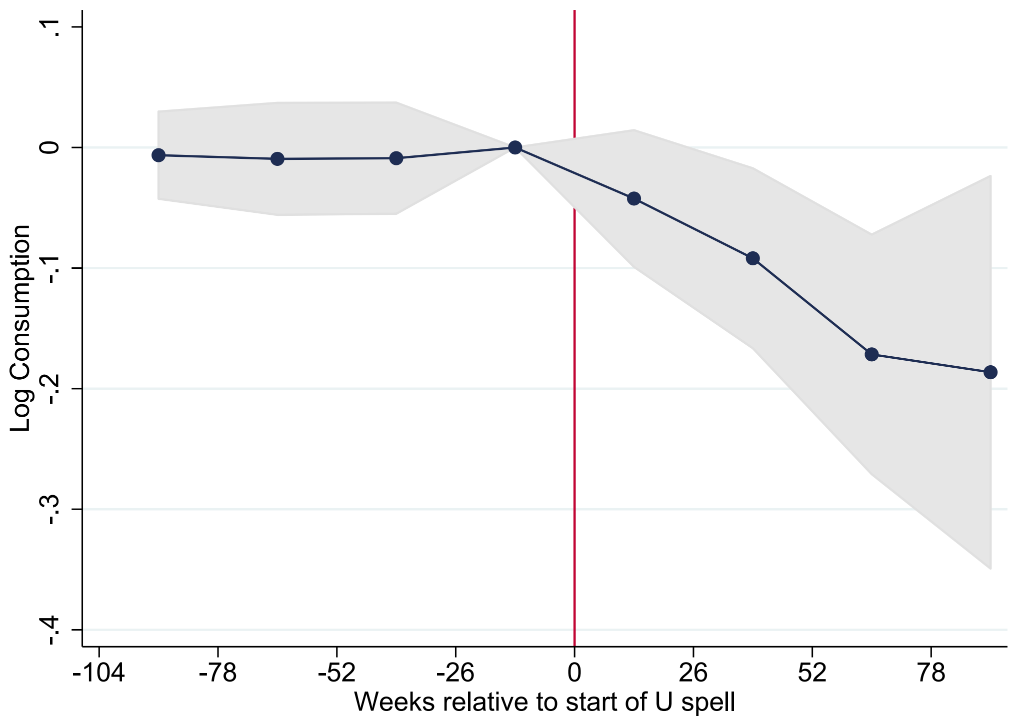
In column (4), we estimate not only differences in levels but also differences in consumption profiles between households who select into spells longer than 20 weeks versus households who select into spells shorter than 20 weeks. To do so, we fully interact $\mathbb{1}[L > 20 \text{ wks}]$ with the event time dummies to estimate two separate consumption profiles over the unemployment spell: one for households with long spells and one for households with short spells. Results reported in column (4) indicate that the consumption drop in the first 20 weeks for households who select into long spells is 1.3% smaller but not statistically different from the consumption drop in the first 20 weeks of households who select into short spells[10]. This evidence suggest that differences in consumption profiles are relatively small and insignificant, so that selection effects are not significantly driving the observed patterns of average household consumption over the unemployment spell.

Tables C-5 and C-6 replicate the analysis of, respectively, column 3 and 4 of Table C-3 for various categories of consumption available in the HUT data. These tables are aimed at gauging how the difference between expenditures and actual consumption may affect our conclusions regarding the evolution of consumption smoothing gains over the unemployment spell, as explained in section 4.2 . The goal is to detect whether individuals who select into long versus short spells have access to different means of smoothing consumption using substitution across different expenditure categories. Results indicate that the expenditure levels and expenditure profiles over food, recreation, transportation, or restaurants are not significantly different for households that select into long versus short spells. This suggests that there is no significant dynamic selection over the spell based on the availability of substitution towards home production. Equivalently, expenditure levels and expenditure profiles over durable goods are not significantly different

[10]To control for length-biased sampling, we also implemented all specifications using reweighting approaches, and with more flexible controls for longer durations L. Results are identical, and available upon request.

for households that select into long versus short spells. The availability of consumption smoothing through shifting expenditures away from durables is not significantly different for households that select into long versus short spells.[11]

Figure C-7: Log household consumption relative to last semester prior to the unemployment spell



Notes: The figure correlates household consumption measured in the HUT survey with the time t since (or until) the onset of the unemployment spell observed in the administrative UI records. Bi-weekly household consumption levels measured at the moment of the HUT interview are annualized and expressed in constant SEK2003. The figure follows from regression model (7) and plots the estimated coefficients β_k for the set of indicators $\mathbb{1}[t \in [26*k, 26*\{k+1\})]$ for being observed in the k-th 6 months period since the onset of one's spell. We also plot the 95% confidence interval from robust standard errors. We include in this regression a set of controls X, which consists of year dummies, calendar months dummies and a set of dummies for family status. The graph provides evidence that average household consumption drops over the unemployment spell. The average consumption of households where a member has been unemployed for a full year is more than 15% lower than the average consumption of household with a member at the start of her spell.

[11]It should be noted that the small sample size in the HUT offers limited statistical power and that it would be interesting to get further evidence on these various expenditure patterns.

Table C-2: Summary statistics at start of unemployment spell: HUT sample

	Mean	P10	P50	P90
	I. Unemployment			
Duration of spell (wks)	26.64	2.86	13.43	65.29
Duration on b_1 (wks)	12.87	2.86	13.43	20
Duration on b_2 (wks)	13.78	0	0	45.29
Replacement rate	.72	.59	.79	.8
	II. Demographics			
Age	34.12	21	33	51
Fraction men	.49	0	0	1
Fraction married	.39	0	0	1
Fraction with higher educ	.26	0	0	1
	III. Income and Wealth, SEK 2003(K)			
Gross earnings (individual)	202.9	9.8	172.6	386.2
Household disposable income	354.4	116.9	330.1	585.3
Household consumption	343	150.3	305.1	572.6
Household net wealth	510.1	-258.3	0	1691.6
Household bank holdings	65.6	0	0	149.8
Household real estate	770.7	0	44	1948.3
Household debt	427.2	0	193.3	1154.3

Notes: The table provides summary statistics for our sample of households unemployed from the HUT consumption surveys 2003 to 2009. To create this sample, we merged the household consumption surveys with the universe of administrative UI records to reconstruct the full employment history of individuals whose household is surveyed in the HUT. We then restrict the sample to households where an individual is either unemployed at the time of the interview, or who will become unemployed some time in the next two years following the interview. This leaves us with a pseudo-panel of about 6,500 households. All earnings, income and asset level measures are from wealth and income registers, and are yearly measures aggregated at the household level in constant k2003SEK for the last calendar year of full employment prior to the start of the unemployment spell. Disposable income is gross earnings, plus capital income minus all taxes plus transfers received. Transfers include unemployment insurance, disability insurance, sick pay, and all housing and parental benefits. All financial assets are estimated at their market value. Real estate is gross of debt and assessed at market value. Debt includes student loans, mortgage, credit card debt, etc. Note that 1SEK2003 $\approx$ 0.11USD2003.

Table C-3: Household log-consumption as a function of time spent unemployed relative to pre-unemployment consumption: consumption survey estimates

	(1)	(2)	(3)	(4)
$\mathbb{1}[0 < t \leq 20 \text{ weeks}]$ (η_1)	-0.0606*	-0.0444	-0.0408	-0.0465
	(0.0316)	(0.0311)	(0.0314)	(0.0413)
$\mathbb{1}[t > 20 \text{ weeks}]$ (η_2)	-0.130***	-0.129***	-0.111***	-0.108***
	(0.0328)	(0.0332)	(0.0386)	(0.0414)
$\mathbb{1}[L > 20 \text{ weeks}]$			-0.0294	-0.0342
			(0.0308)	(0.0378)
$\mathbb{1}[0 < t \leq 20 \text{ weeks}] \times \mathbb{1}[L > 20 \text{ weeks}]$				0.0134
				(0.0629)
Year F-E	×	×	×	×
Calendar months F-E	×	×	×	×
Marital status		×	×	×
Family size		×	×	×
Age group F-E		×	×	×
Test $\eta_1 = \eta_2$				
P-value	**0.06**	**0.02**	**0.08**	**0.10**
R^2	0.0493	0.0866	0.0872	0.0872
N	1551	1548	1548	1548

Notes: Robust standard errors in parentheses. * p<.10, ** p<.05, *** p<.01.
The Table reports estimates of the drop in household consumption over the unemployment spell in the HUT surveys, following model of equation (8). Because the HUT surveys collect information on household consumption expenditures at the time of the interview, HUT estimates can directly recover flow (bi-weekly) measures of consumption c_t. We restrict the sample to households where, at the date of the interview, one (and only one) individual is unemployed, or where, at the date of the interview, one (and only one) individual will become unemployed in the following 6 months. $\mathbb{1}[0 < t <= 20 \text{ weeks}]$ is an indicator for having a member of the household unemployed since less than 20 weeks at the time of the interview. $\mathbb{1}[t > 20 \text{ weeks}]$ is an indicator for having a member of the household unemployed for more than 20 weeks at the time of the interview. $\mathbb{1}[L > 20 \text{ weeks}]$ is an indicator for the total duration of the unemployment spell being longer than 20 weeks. We also report the P-value of a test of equality of η_1 (the drop in consumption in the first 20 wks) and η_2 (the drop in consumption after 20 wks).

Table C-4: Consumption as a function of time spent unemployed relative to pre-unemployment consumption: consumption survey estimates for various expenditure categories

	(1) Total expenditures	(2) Food	(3) Rents	(4) Purchase of new vehicles	(5) Furniture & house appliances	(6) Trans-portation	(7) Recre-ation	(8) Restau-rant
$\mathbb{1}[0 < t \leq 20$ weeks$]$	-0.0606*	-0.0441	-0.0404	-0.418**	-0.160	-0.0788	-0.106	-0.0807
	(0.0316)	(0.0388)	(0.0380)	(0.187)	(0.102)	(0.0661)	(0.0649)	(0.0876)
$\mathbb{1}[t > 20$ weeks$]$	-0.130***	-0.0823*	0.0430	-0.252	-0.0883	-0.348***	-0.189***	-0.165*
	(0.0328)	(0.0441)	(0.0310)	(0.176)	(0.0884)	(0.0803)	(0.0719)	(0.0888)
Year fixed effects	×	×	×	×	×	×	×	×
Calendar months F- E	×	×	×	×	×	×	×	×
R^2	0.0493	0.0650	0.0365	0.0205	0.00975	0.0208	0.0252	0.0154
N	1551	1548	798	982	1548	1488	1543	1119

Notes: Robust standard errors in parentheses. * p<.10, ** p<.05, *** p<.01.
The Table reports estimates of the drop in household consumption over the unemployment spell in the HUT surveys, following model of equation (8). Because the HUT surveys collect information on household consumption expenditures at the time of the interview, HUT estimates can directly recover flow (bi-weekly) measures of consumption c_t. We restrict the sample to households where, at the date of the interview, one (and only one) individual is unemployed, or where, at the date of the interview, one (and only one) individual will become unemployed in the following 6 months. $\mathbb{1}[0 < t <= 20$ weeks$]$ is an indicator for having a member of the household unemployed since less than 20 weeks at the time of the interview. $\mathbb{1}[t > 20$ weeks$]$ is an indicator for having a member of the household unemployed for more than 20 weeks at the time of the interview. $\mathbb{1}[L > 20$ weeks$]$ is an indicator for the total duration of the unemployment spell being longer than 20 weeks.

Table C-5: LOG-CONSUMPTION AS A FUNCTION OF TIME SPENT UNEMPLOYED RELATIVE TO PRE-UNEMPLOYMENT CONSUMPTION: CONSUMPTION SURVEY ESTIMATES FOR VARIOUS EXPENDITURE CATEGORIES, SELECTION ON LEVELS

	(1) Total expenditures	(2) Food	(3) Rents	(4) Purchase of new vehicles	(5) Furniture & house appliances	(6) Trans-portation	(7) Recre-ation	(8) Restau-rant
$\mathbb{1}[0 < t \leq 20$ weeks$]$	-0.0379	-0.0139	-0.0212	-0.192	-0.0657	-0.0596	-0.0948	-0.0765
	(0.0305)	(0.0366)	(0.0368)	(0.172)	(0.0893)	(0.0658)	(0.0649)	(0.0876)
$\mathbb{1}[t > 20$ weeks$]$	-0.113***	-0.0782*	0.0143	-0.387*	-0.0463	-0.330***	-0.146*	-0.141
	(0.0379)	(0.0463)	(0.0356)	(0.217)	(0.102)	(0.0925)	(0.0850)	(0.104)
$\mathbb{1}[L > 20$ weeks$]$	-0.0294	-0.00917	0.0239	0.0106	-0.0394	-0.0223	-0.0775	-0.0208
	(0.0300)	(0.0364)	(0.0312)	(0.159)	(0.0848)	(0.0659)	(0.0654)	(0.0811)
Year fixed effects	×	×	×	×	×	×	×	×
Marital status	×	×	×	×	×	×	×	×
Family size	×	×	×	×	×	×	×	×
R^2	0.139	0.194	0.157	0.0686	0.0442	0.0525	0.0567	0.0369
N	1548	1545	796	606	1523	1485	1540	1116

Notes: Robust standard errors in parentheses. * p<.10, ** p<.05, *** p<.01.
The Table reports estimates of the drop in household consumption over the unemployment spell in the HUT surveys, following model of equation (8). Because the HUT surveys collect information on household consumption expenditures at the time of the interview, HUT estimates can directly recover flow (bi-weekly) measures of consumption c_t. We restrict the sample to households where, at the date of the interview, one (and only one) individual is unemployed, or where, at the date of the interview, one (and only one) individual will become unemployed in the following 6 months. $\mathbb{1}[0 < t <= 20$ weeks$]$ is an indicator for having a member of the household unemployed since less than 20 weeks at the time of the interview. $\mathbb{1}[t > 20$ weeks$]$ is an indicator for having a member of the household unemployed for more than 20 weeks at the time of the interview. $\mathbb{1}[L > 20$ weeks$]$ is an indicator for the total duration of the unemployment spell being longer than 20 weeks.

Table C-6: Log-consumption as a function of time spent unemployed relative to pre-unemployment consumption: consumption survey estimates for various expenditure categories, selection on profile

	(1) Total expenditures	(2) Food	(3) Rents	(4) Purchase of new vehicles	(5) Furniture & house appliances	(6) Trans-portation	(7) Recre-ation	(8) Restau-rant
$\mathbb{1}[0 < t \leq 20$ weeks$]$	-0.0465	-0.0320	-0.00595	-0.288	0.0324	-0.0457	-0.126	-0.107
	(0.0413)	(0.0527)	(0.0456)	(0.223)	(0.114)	(0.0920)	(0.0879)	(0.112)
$\mathbb{1}[t > 20$ weeks$]$	-0.108***	-0.0666	0.00371	-0.347	-0.0976	-0.328***	-0.127	-0.117
	(0.0414)	(0.0528)	(0.0442)	(0.218)	(0.116)	(0.0935)	(0.0881)	(0.115)
$\mathbb{1}[L > 20$ weeks$]$	-0.0342	-0.0258	0.0440	-0.0582	0.0441	-0.0147	-0.0967	-0.0410
	(0.0378)	(0.0482)	(0.0409)	(0.193)	(0.105)	(0.0848)	(0.0804)	(0.102)
$\mathbb{1}[L > 20$ weeks$] \times \mathbb{1}[0 < t \leq 20$ weeks$]$	0.0134	0.0265	-0.0509	0.022	-0.024	-0.0239	0.0750	0.0990
	(0.0629)	(0.0800)	(0.0691)	(0.342)	(0.175)	(0.141)	(0.134)	(0.172)
Year fixed effects	×	×	×	×	×	×	×	×
Marital status	×	×	×	×	×	×	×	×
Family size	×	×	×	×	×	×	×	×
R^2	0.0872	0.107	0.0765	0.0689	0.0273	0.0373	0.0348	0.0246
N	1548	1545	796	606	1523	1485	1540	1116

Notes: Robust standard errors in parentheses. * p<.10, ** p<.05, *** p<.01.
The Table reports estimates of the drop in household consumption over the unemployment spell in the HUT surveys, following model of equation (8). Because the HUT surveys collect information on household consumption expenditures at the time of the interview, HUT estimates can directly recover flow (bi-weekly) measures of consumption c_t. We restrict the sample to households where, at the date of the interview, one (and only one) individual is unemployed, or where, at the date of the interview, one (and only one) individual will become unemployed in the following 6 months. $\mathbb{1}[0 < t <= 20$ weeks$]$ is an indicator for having a member of the household unemployed since less than 20 weeks at the time of the interview. $\mathbb{1}[t > 20$ weeks$]$ is an indicator for having a member of the household unemployed for more than 20 weeks at the time of the interview. $\mathbb{1}[L > 20$ weeks$]$ is an indicator for the total duration of the unemployment spell being longer than 20 weeks.

FOR ONLINE PUBLICATION - Appendix D: Structural Welfare Analysis

This Appendix provides more detail on the specification, approach and results of our calibration exercise.

D.1 Calibration

Our structural model builds on the search environment considered in Hopenhayn and Nicolini [1997], Lentz and Traneas [2005] and Chetty [2008a]. We consider agents with CRRA preferences with additive, iso-elastic search costs,

$$\begin{aligned} v_{i,t}^{u}\left(c_{i,t}, s_{i,t}\right) &= c_{i,t}^{1-\gamma} / \left(1-\gamma\right) - \psi_0 \left(s_{i,t}\right)^{\psi_1}, \\ v_{i,t}^{e}\left(c_{i,t}\right) &= c_{i,t}^{1-\gamma} / \left(1-\gamma\right). \end{aligned}$$

In contrast with the previous models, we allow for individual-specific, duration-dependent exit rate functions

$$h_{i,t}\left(s_{i,t}\right) = h_0 + \left[1 + \exp\left(-\theta t\right)\right] \kappa_i h_1 \times s_{i,t}^{\rho},$$

where h_1 determines the return to search, θ is the exponential rate at which these returns depreciate, κ_i is an individual-specific scalar, drawn from a gamma distribution $\Gamma\left(a,b\right)$, and h_0 is a baseline exit rate (at zero effort).

Each household starts the unemployment spell with an asset level $a_{i,1}$, drawn from a Singh-Maddala distribution $F(a|\alpha_{SM}, c_{SM}, k_{SM})$ (Singh and Maddala [1976]). Each household can draw down their asset to increase consumption as long as $a_{i,t} \geq \bar{a}$, where $\bar{a}$ is a uniform asset limit. We choose the distribution parameters to match the average, 25-th and 90-th percentile of the household distribution of liquid assets.

We consider a flat benefit policy that replaces 72% of the pre-unemployment earnings, normalized to 1. When employed, the agent pays a (uniform) tax τ equal to 5%, which balances the expected revenues and expenditures for the average Swedish unemployment rate between 1999 and 2006 of .069. Our calibration targets household consumption levels, we therefore augment the individual's income with a household income of .7, which corresponds to the average share of other household members' earnings relative to the earnings of the unemployed in our sample (see Table 2).

We set the parameter of CRRA $\gamma = 2$, discount factor $\beta = 1$ and interest rate $r = 0$. All other parameters are used to calibrate the model. We solve the model using an asset grid of 500 grid points and 10 different deciles of the distribution of individual-specific returns to search. For each decile we solve the model backwards starting in week $T = 520$ and then simulate the economy forward for each asset level. We then aggregate up the time paths for consumption and the survival rates using the joint distribution over assets and individual-specific returns to search. We compute the elasticities and average durations implied by the model.

D.2 Model Fit

Table D-1 lists the targets of the calibration, which are the variables underlying our sufficient statistics. The targets include the benefit duration elasticities $\varepsilon_{D,b_{\mathbf{k}}}$, average benefit durations $D_{\mathbf{k}}$ and the average consumption drops $\left[\bar{c}_{\mathbf{k}} - \bar{c}_0\right]/\bar{c}_0$. The benchmark consumption level $\bar{c}_0$ is the average level of consumption when agents would have been employed from the start of the model. For each of the target variables, we target the level for the first part of the policy, as well as the ratio for the two parts and we put higher weight on the latter. Our minimization routine searches for the vector of parameter values that minimizes the relative difference between the targeted moments and the simulated values of these moments generated by our model and is robust to different starting values for the parameters. Table D-1 compares the targeted and estimates values of these moments. The calibration slightly overestimates the consumption drop in the first 20 weeks and underestimates the consumption drop thereafter. The simulated drops are well within the 95% confidence intervals of our empirical estimates, but imply that we underestimate the $CS_{\mathbf{2}}/CS_{\mathbf{1}}$ compared to our sufficient-statistics implementation. Our calibrated model also underestimates the unemployment durations and elasticities, but closely matches the relative moral hazard costs $MH_{\mathbf{2}}/MH_{\mathbf{1}}$. The model underestimates the unemployment durations and elasticities, but matches the ratios very well. The calibrated parameters for the returns to search imply both strong depreciation of the exit rates and heterogeneity in exit rates. The average exit rate equals .21 in the first week of the spell and .03 in the 20-th week of unemployment, while the exit rate of a given job seeker is about half as large in the 20-th week compared to the first week. The model also predicts that 53% of the unemployed are liquidity constrained and consume hand-to-mouth from the start of the unemployment spell.

Table D-1: Model Calibration: Targeted and Estimated values

Targets	Description	Targeted Value	Estimated Value
$a_{.25}$	25th percentile of asset distribution	0.00	0.000
$a_{.90}$	90th percentile of asset distribution	0.42	0.553
a_{mean}	mean of asset distribution	0.19	0.203
$\Delta c_{\mathbf{1}}$	consumption drop ST unemployed	0.060	0.072
$\Delta c_{\mathbf{1}}/\Delta c_{\mathbf{2}}$	ratio of consumption drops	0.462	0.658
$D_{\mathbf{1}}$	ST benefit duration	12.600	6.675
$D_{\mathbf{1}}/D_{\mathbf{2}}$	ratio of benefit duration	0.940	0.991
$\varepsilon_{D_{b_{\mathbf{1}}}}$	elasticity wrt ST benefits	0.854	0.549
$\varepsilon_{D_{b_{\mathbf{1}}}}/\varepsilon_{D_{b_{\mathbf{2}}}}$	ratio of elasticities	1.262	1.271

Notes: The table lists the variables we target in our calibration exercise. These are the moments underlying our sufficient statistics such that the structural model provides the same policy recommendations as implied by the sufficient-statistics approach. For each of the target variables, we target the level for the first part of the policy, as well as the ratio for the two parts. The parameters of the asset distribution are set to match the mean, 25-th and 90-th percentile of the household distribution of liquid assets (expressed relative to annual household income). We also set the parameter of CRRA $\gamma = 2$, discount factor $\beta = 1$ and interest rate $r = 0$. The remaining model parameters are calibrated to minimize the relative distance between the targeted moments and the simulated values of these moments generated by our model.

D.3 Counterfactual Analysis

The calibrated model allows analyzing the moral hazard costs and consumption smoothing gains for different unemployment policies, keeping all parameters constant, including the asset distribution at the start of the unemployment spell. The government's budget is affected by the changes in the benefit levels (as we keep the tax rate fixed), but these budgetary changes are traded off against the agents' welfare, using the normalization of the Lagrange multiplier on the government's constraint, $\lambda \equiv E_0 \left(v' \left(c_{i,0} \right) \right)$, where $c_{i,0}$ is the agent's level of consumption when she would have been employed from the start of the model. Figure 7 plots the respective consumption smoothing gains and moral hazard costs. $CS_{\bar{b}}$ and $MH_{\bar{b}}$ have a unique intersection for $\bar{b} = .58$, which corresponds to the optimal flat benefit level for the given tax rate and Lagrange multiplier. Similarly, we find that $CS_{\mathbf{1}} = MH_{\mathbf{1}}$ and $CS_{\mathbf{2}} = MH_{\mathbf{2}}$ jointly hold for the unique pair $(b_{\mathbf{1}}, b_{\mathbf{2}}) = (.48, .68)$.

We repeat the previous analysis for two variations of our baseline model. In the model with only heterogeneity in the returns to search, we set the search depreciation parameter $\theta = 0$, but re-calibrate the return to search parameter h_1 to maintain the same average duration. All other parameters remain the same. In the model with only depreciation of the returns to search, we assume a uniform scalar $\kappa_i = \bar{\kappa}$ and re-calibrate h_1 to maintain the same average duration. Figure D-1 illustrates the respective consumption smoothing gains and moral hazard costs. The simulated implications for the flat profile are similar, except that the unemployment elasticities and thus the moral hazard costs are higher in the model with only depreciation (but no heterogeneity). In both alternative models, $MH_{\bar{b}}$ decreases while $CS_{\bar{b}}$ increases when the replacement rate is reduced. They are equalized for $\bar{b} = .60$ in the model with only heterogeneity and $\bar{b} = .40$ in the model with only depreciation. Regarding the tilt, the model with only depreciation of exit rates increases $MH_{\mathbf{2}}$ relative to $MH_{\mathbf{1}}$.[12] This calibration pushes towards a declining tilt and more so as the overall generosity is reduced. However, the model with only heterogeneity continues to prescribe an inclining tilt and this remains true for lower replacement rates. Importantly, conditional on the values of the sufficient statistics, the underlying depreciation and heterogeneity do not affect the local policy recommendations. The differential impact on how the sufficient statistics change with the unemployment policy, however indicates the complementary value of a structural approach and disentangling the two non-stationary forces to go beyond our local recommendations.

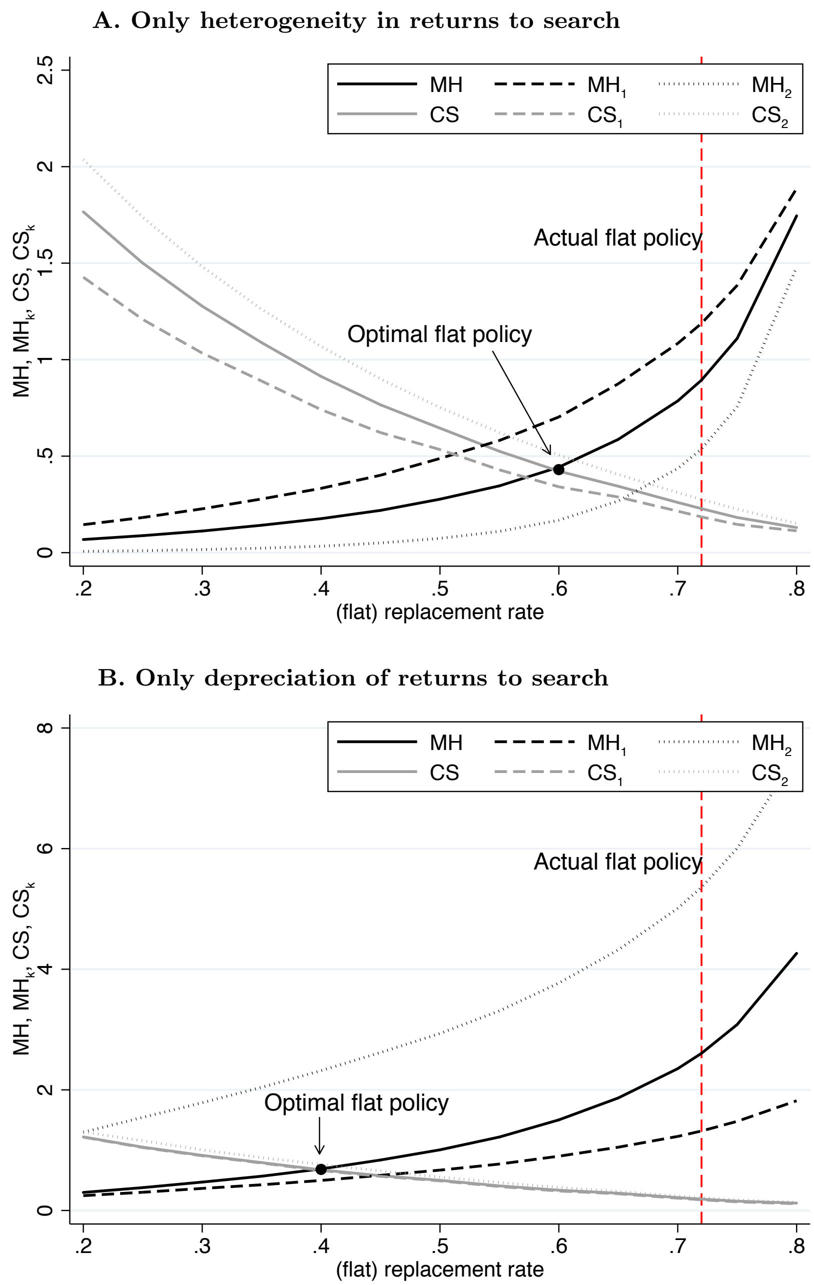
We also note that the structural model can be exploited to assess the potential importance of some implementation assumptions. We simulate the consumption drops and calculate the implied consumption smoothing gains for different levels of risk aversion, reported in Table D-2, while keeping all other parameters constant. The Taylor approximation of the marginal CRRA utilities tends to underestimate the consumption smoothing gains, especially when the level of risk aversion is high. This was already noted in Chetty [2006]. Interestingly, the approximation error on the relative consumption smoothing gains is very small and thus the recommendation on the tilt unaffected, regardless of the level of risk aversion.

We finally use the structural model to gauge the impact of heterogeneity in preferences on the estimated consumption smoothing gains. We simulate a model with share α of job seekers with risk aversion $\gamma = 1$ and share $1 - \alpha$ with risk aversion $\gamma = 3$. We set $\alpha = .58$ such that the average risk aversion among the unemployed, accounting for type-specific unemployment durations, equals $\gamma = 2$. Since the consumption drops hardly change for agents with different preferences, the average consumption drops in the model with heterogeneity remain very similar compared to the baseline model. The dynamic selection on preferences could in principle affect the relative consumption smoothing gains, as shown in subsection 5.2.3. Note that in our model with CRRA preferences and additive search cost (a standard specification also used in Hopenhayn and Nicolini [1997] and Chetty [2008a]), individuals with higher risk aversion have lower marginal utility of consumption (relative to the disutility of search), therefore search less and leave unemployment more slowly. As illustrated in equation (21) in subsection 1.3, dynamic selection on both the marginal utility of consumption $E_k \left[v_i' \left(c_{i,0} \right) \right] / E_0 \left[v_i' \left(c_{i,0} \right) \right]$ and the consumption smoothing gains $E_k \left[v_i'' \left(c_{i,0} \right) \left(c_{i,0} - c_{i,t}^u \right) \right] / E_0 \left[v_i' \left(c_{i,0} \right) \right]$ affect the value of CS_k. Hence, the net impact is ambiguous. In our simulation with heterogeneity, the aggregate consumption smoothing gains $CS_{\mathbf{1}}$ and $CS_{\mathbf{2}}$ are lower relative to the baseline

[12] Note that in Appendix C we demonstrated that depreciation in the search returns can increase $MH_{\mathbf{1}}$ relative to $MH_{\mathbf{2}}$. This comparative static was derived without keeping the average duration constant.

scenario. The ratio $CS_{\mathbf{1}}/CS_{\mathbf{2}}$ decreases as well, which further increases the value of introducing an inclining tilt ($b_{\mathbf{2}} > b_{\mathbf{1}}$). We note though that with heterogeneous preferences, the calculation of $CS_{\mathbf{k}}$ critically depends on the normalization $\lambda \equiv E_0\left(v'\left(c_{i,0}\right)\right)$. We can avoid this normalisation by directly considering the relative consumption smoothing gains $\left[1 + CS_{\mathbf{1}}\right] / \left[1 + CS_{\mathbf{2}}\right] = E_1\left(v_i'\left(c_{i,t}\right)\right) / E_2\left(v_i'\left(c_{i,t}\right)\right)$, relevant for evaluating a budget-balanced change in the tilt. This ratio changes from .934 in the baseline calibration to .935 in the simulation with heterogeneity (cf. Table D-2) Overall, the introduction of preference heterogeneity seems to limited impact on the relative consumption smoothing gains in our simulations and, if anything, increases the welfare gain from introducing an inclining tilt ($b_{\mathbf{2}} > b_{\mathbf{1}}$).

Figure D-1: Alternative Model Specifications: Welfare Effects for Different Benefit Levels



Notes: The figure analyzes the robustness of the findings in Figure 7 for different model specifications. The two panels plot the moral hazard costs and consumption smoothing gains for different levels of the flat benefit profile. Panel A uses a model without depreciation of the returns to search and only heterogeneity in the returns to search. Panel B uses a model with only depreciation of the returns to search and no heterogeneity in the returns to search. Compared to our baseline model, we re-calibrate the uniform return to search parameter to maintain the same average exit rate. The actual policy is a flat profile with average replacement rate of .72 and is indicated by the vertical dashed line. We report the simulated moral hazard costs and consumption smoothing gains for an overall change in the flat benefit profile $\bar{b}$, for an increase in the benefit level in the first 20 weeks of unemployment, and for an increase in the benefit level after 20 weeks of unemployment.

Table D-2: Implementation of consumption smoothing gains

		Consumption Smoothing Gain		
	$\Delta c_k/c$	$E_k\left[v_i'\left(c_{i,t}\right)\right]$	CS_k	$\gamma\Delta c_k/c$
	Baseline Calibration: γ=2			
Throughout spell	0.091	0.443	0.231	0.181
First 20 weeks	0.072	0.428	0.189	0.143
After 20 weeks	0.109	0.458	0.273	0.218
Before 20wks / after 20 wks	0.658	0.934	0.693	0.658
	Simulation with heterogenous preferences			
Throughout spell	0.091	0.485	0.108	0.182
First 20 weeks	0.072	0.469	0.070	0.144
After 20 weeks	0.109	0.501	0.144	0.217
Before 20wks / after 20 wks	0.664	0.935	0.483	0.664
	Simulation with low risk aversion γ=1			
Throughout spell	0.093	0.666	0.109	0.093
First 20 weeks	0.074	0.653	0.088	0.074
After 20 weeks	0.112	0.677	0.129	0.112
Before 20wks / after 20 wks	0.658	0.964	0.682	0.658
	Simulation with high risk aversion γ=3			
Throughout spell	0.088	0.296	0.368	0.264
First 20 weeks	0.071	0.282	0.307	0.212
After 20 weeks	0.105	0.309	0.429	0.316
Before 20wks / after 20 wks	0.671	0.914	0.715	0.671

Notes: This Table illustrates alternative implementations of the consumption smoothing gains for different underlying risk preferences. The top panel reports the simulated consumption drops and consumption smoothing gains for the baseline model with homogeneous CRRA γ=2. The third and fourth panel report the simulated values for lower CRRA γ=1 and higher CRRA γ=3, but still assuming homogenous preferences. The second panel reports the simulated values for a model with a mixture of these two preference types, keeping the average CRRA among the unemployed equal to 2. This makes the results comparable to the baseline model. The first column in each panel shows the drops in consumption during the respective parts of the unemployment spell relative to the consumption levels when employed at the start of the model. The second column shows the average marginal utility of consumption during the respective parts of the unemployment spell. The third column reports $CS_{\mathbf{k}}$, which expresses the average marginal utility of consumption in the second column relative to the average marginal utility of employment consumption at the start of the model. The fourth column provides an approximation of $CS_{\mathbf{k}}$ following the implementation in (14), which relies on a Taylor approximation and homogeneous preferences. For this fourth column, the second panel with heterogeneous preferences simply multiplies the consumption drops by γ=2. The rows of each panel report the consumption drops and consumption smoothing gains over the full unemployment spell, during the first 20 weeks of unemployment and after 20 weeks of unemployment respectively. The last row shows the ratio of the first part to the second part of the unemployment spell.

References -Appendix A to Appendix D

Blanchard, Olivier Jean and Peter Diamond, "Ranking, Unemployment Duration, and Wages," *The Review of Economic Studies*, 1994, *61* (3), 417–434.

Browning, Martin and Søren Leth-Petersen, "Imputing consumption from income and wealth information," *The Economic Journal*, 2003, *113* (488), F282–F301.

Chetty, Raj, "A General Formula for the Optimal Level of Social Insurance," *Journal of Public Economics*, 2006, *90* (10-11), 1879–1901.

— , "Moral Hazard versus Liquidity and Optimal Unemployment Insurance," *Journal of Political Economy*, 2008, *116* (2), 173–234.

Cohen, Alma and Liran Einav, "Estimating Risk Preferences from Deductible Choice," *American Economic Review*, 2007, *97* (3), 745–788.

Eika, L., M. Mogstad, and O. Vestad, ""What Can We Learn About Household Consumption From Information on Income and Wealth"," *working paper*, 2017.

Ganong, Peter and Simon Jaeger, "A Permutation Test and Estimation Alternatives for the Regression Kink Design," *Harvard University Discussion Paper*, 2014.

Hopenhayn, Hugo A. and Juan Pablo Nicolini, "Optimal Unemployment Insurance," *Journal of Political Economy*, 1997, *105* (2), 412–438.

Koijen, Ralph, Stijn Van Nieuwerburgh, and Roine Vestman, *Judging the Quality of Survey Data by Comparison with Truth as Measured by Administrative Records: Evidence From Sweden*, University of Chicago Press, July 2014

Kolsrud, Jonas, Camille Landais, and Johannes Spinnewijn, "Registry-Based Measures of Consumption And Consumption Patterns: What Can We Learn From the Swedish Administrative Data?," *Mimeo*, 2017.

Kroft, Kory and Matthew J Notowidigdo, "Should unemployment insurance vary with the unemployment rate? Theory and evidence," *The Review of Economic Studies*, 2016, p. rdw009.

Landais, Camille, "Assessing the Welfare Effects of Unemployment Benefits Using the Regression Kink Design," *American Economic Journal: Economic Policy*, 2015.

— , Pascal Michaillat, and Emmanuel Saez, "Optimal Unemployment Insurance over the Business Cycle," Working Paper 16526, National Bureau of Economic Research 2010.

Lehr, Brandon, "Optimal Unemployment Insurance with Endogenous Negative Duration Dependence," *Public Finance Review*, 2017, *45* (3), 395–422.

Lentz, Rasmus and Torben Traneas, "Job Search and Savings: Wealth Effects and Duration Dependence," *Journal of Labor Economics*, July 2005, *23* (3), 467–490.

Lockwood, Ben, "Information Externalities in the Labour Market and the Duration of Unemployment," *The Review of Economic Studies*, 1991, *58* (4), 733–753.

Meyer, Bruce, "Unemployment Insurance and Unemployment Spells," *Econometrica*, 1990, *58(4)*, 757–782.

Michaillat, Pascal, "Do Matching Frictions Explain Unemployment? Not in Bad Times," *American Economic Review*, 2012, *102* (4), 1721–50.

Shavell, Steven and Laurence Weiss, "The Optimal Payment of Unemployment Insurance Benefits over Time," *Journal of Political Economy*, 1979, *87* (6), 1347–1362.

Singh, S. and G. Maddala, "A Function for the Size Distribution of Incomes," *Econometrica*, 1976, *44* (5), 963–970.

Spinnewijn, Johannes, "Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs," *Journal of the European Economic Association*, 2015, *13* (1), 130–167.