The Trade-off between Insurance and Incentives in Differentiated Unemployment Policies

Johannes Spinnewijn
London School of Economics

August 9, 2019

Abstract

The article revisits the central trade-off between insurance and incentives in the design of UI policies. The generosity of UI benefits does not just differ across countries, but also across workers within countries. After illustrating some important dimensions of heterogeneity in a cross-country analysis, I extend the standard Baily-Chetty formula to identify the key empirical moments and elasticities required to evaluate the differentiated unemployment policy within a country. The article reviews some prior work and aims to provide guidance for future work trying to inform the design of unemployment policies.

Keywords: Unemployment, Policy Design, Heterogeneity

JEL codes: H20, J64, J65

1 Introduction

One of the Institute for Fiscal Studies’ central adages is to bring together theoretical models and empirical analysis to inform policy. Over the past fifty years, the research on the design of unemployment insurance has slowly adapted this adage, which has resulted in a rich literature that tightly integrates theory and empirics and allows for comprehensive, evidence-based policy evaluations.

It was at the end of the first decade after the IFS was founded that Martin Baily characterized the central trade-off for unemployment insurance design in his paper “Some Aspects of Optimal Unemployment Insurance” published in the Journal of Public Economics in 1978. The trade-off is to provide insurance against the income

---

*This article has immensely benefited from my collaborations on UI design with various co-authors and with my colleague Camille Landais in particular. I would also like to thank Gabriel Leite Mariani and Tatiana Pazem for excellent research assistance, and the editor Monica Costa Dias and two anonymous referees for valuable comments and suggestions. I acknowledge financial support from the ERC starting grant #716485.
loss due to unemployment, while maintaining incentives to reduce unemployment risk. In this classic paper Baily emphasized the value that social insurance can provide, in response to emerging work that warned for negative consequences of UI through lower employment incentives and increased the unemployment rates (Feldstein 1973, 1974). He expressed the trade-off in a simple formula consisting of few moments, capturing the consumption smoothing gains and the unemployment costs of UI respectively. These moments can be readily taken to the data.

It took another twenty years until the early days of the empirical “credibility revolution” when Jonathan Gruber was the first to implement the Baily formula empirically (Gruber 1997). This spurred a significant amount of follow-up work, which took off further when Raj Chetty demonstrated the theoretical credibility of Baily’s simple formula ten years later (Chetty 2006). Chetty showed that the moments of the formula remain the same when enriching the underlying theoretical model. These moments are often referred to as “sufficient statistics” for the evaluation of UI policy (Chetty 2009), akin to the elasticity of taxable income for tax policy.

Today, at the fiftieth anniversary of the IFS, the body of research on UI and the evaluation of its design in particular is impressive (see Chetty and Finkelstein 2013, Tatsiramos and Van Ours 2014 and Schmieder and Von Wachter 2016 for reviews). Also Baily’s formula has been extended and implemented to account for various other features that take us closer to the real world. This ranges from accounting for general equilibrium effects (Landais et al. 2018a,b) and behavioral frictions (Spinnewijn 2015), to including informal sector work (Gerard and Gonzaga 2016), allowing for duration-dependence (Nekoei and Weber 2017) and for dynamic selection (Kolsrud et al. 2018). The work on UI relative to other social insurance programs is arguably disproportional to the importance of UI policy as measured by their share of government expenditures. However, our improved understanding of the trade-off between insurance and incentives is as applicable to other social insurance programs, like health insurance and social security. The focus on UI is partly explained by the presence of tractable workhorse model (which is arguably lacking for other social insurance programs) and the remarkable variation in UI policy and rules, allowing for the estimation of its causal effects.

Somewhat surprisingly, the evaluation of the specific policy rules used in the design of UI and resulting differentiation in UI generosity has been lagging behind. In fact, Martin Baily already acknowledged this omission in his original article in 1978, stating: “A compulsory government program prevents adverse selection from driving out the insurance coverage, but of course it is still true that when workers are not all alike, some of them have much more to gain from the program than others, and I am ignoring this.” A number of studies have contributed to filling this gap: e.g., Shavell and Weiss (1979); Hopenhayn and Nicolini (1997); Shimer and Werning (2008); Kolsrud et al. (2018); Lindner and Reizer (2018) on the dynamic profile of benefits, Schmieder et al. (2013).
Kroft et al. (2013); Landais et al. (2018a,b) on benefits over the business cycle, Michelacci and Ruffo (2015) on benefits over the life-cycle. But, overall, understanding the vast differences in UI policies both across and within countries is an understudied area.

The question how to optimally differentiate the unemployment policy - either based on exogenous conditions or on workers’ choices - closely relates to the question on which tags to use to target social transfers and to improve redistribution (Akerlof, 1978). In this article I will argue that the Baily formula and the sufficient statistics insights can be valuable to think about the design of differentiated UI. The main objective of this paper is thus to demonstrate the policy-relevance of further research filling this gap in the UI literature and to provide some guidance on how to get started.

The first part of the paper illustrates the variation in UI policy rules across countries and the within-country variation in UI generosity the policy rules entail. I draw some key lessons from this institutional analysis. First, the main dimensions along which the UI generosity differs across different countries are pre-unemployment earnings and the length of the unemployment spell. Second, beyond these two dimensions, there are various other rules and conditions that countries apply. The desirability of these features seems untested, but the amalgamation of rules increases the complexity for potential beneficiaries, which seems undesirable in itself. Finally, as the variation in UI generosity within countries is substantial, a useful cross-country comparison of overall UI generosity is difficult. Moreover, a cross-country evaluation of UI generosity is complicated by the different context and objectives that apply in different countries, which is less of a constraint for a within-country evaluation of the differentiated UI policies.

The second part of the paper revisits the trade-off between insurance and incentives as characterized by Baily (1978), but extending the framework to evaluate the optimal differentiation of UI policy. This extension builds on Kolsrud et al. (2018) who study how unemployment benefits should evolve over the unemployment spells and allows for some simple insights. First and foremost, the respective consumption smoothing gains and unemployment responses for the groups subject to different UI generosity are key (and sometimes sufficient) to evaluate the desirability of the differentiation itself. This highlights the value of further empirical work estimating heterogeneity in the insurance gains and incentives, in particular along the particular dimensions used for the differentiation of UI benefits. The challenge is that the within-group policy variation required to estimate the relevant moments may not be readily available. Second, the evaluation is complicated when the eligibility rules are based on workers’ choices, either directly or indirectly. Examples are when unemployment benefits are conditional on the time spent unemployed or the age at the onset of the unemployment spell. In general, one needs to estimate the response in the share of workers eligible to any benefit level to the specific change in the UI policy. Conveniently, these responses are of second-order importance when we start from a uniform policy and simply consider
in which direction to differentiate the policy.

I conclude this article by discussing the empirical implementation of the insurance-incentive trade-off and highlight some of the caveats. In particular, the recommendations on the optimal differentiation across groups can be more robust and/or less data-demanding than the evaluation of the optimal level for any given group. I also open the discussion on whether the analysis can be extended to start thinking beyond the traditional UI programs that condition benefit receipt on unemployment as to provide better protection against job loss in today’s labor markets. Throughout my discussion, I will review some recent contributions to the literature, but I again refer the interested reader to the aforementioned reviews for more comprehensive accounts of the relevant literature.

2 UI Design in Practice

This section illustrates the substantial variation in UI generosity not only across, but also within countries. Since countries differentiate their UI policies in very different ways, the protection from which a worker can benefit when unemployed is very different depending on the country and the circumstances he or she is in. My focus is primarily on standard UI benefits, designed to (partially) replace pre-unemployment earnings while unemployed. In many countries, unemployed workers who are not entitled to (or exhausted) standard UI benefits can often rely on what the OECD defines as Unemployment Assistance (UA). These transfers are in general less generous and means-tested. I account for these transfers in the calculations below. Individuals who lose their job or out of work may also benefit from other forms of income support or welfare (e.g., housing benefits, child benefits, food stamps, basic income, etc), which are ignored in the calculations. The UI benefits, conditional on unemployment, can also be complemented with government-mandated severance payments, conditional on layoff. Severance payments have become more important over time and are the primary source of insurance against job loss in developing countries (Gerard and Nairitomi 2019), but are also ignored in the calculations below.

I have studied the UI policy rules for 18 countries, which include the UK, the US and Canada, and a large number of European countries for which the data availability and UI policy variation has been used in recent papers. This includes Austria (Card et al. 2007; Nekoei and Weber 2017), Germany (Schmieder et al. 2016), Hungary (DellaVigna et al. 2017; Lindner and Reizer 2018), Italy (Citino et al. 2019), and Sweden (Kolsrud et al. 2018; Landais et al. 2017; Landais and Spinnewijn 2018). I also include Japan and South-Korea. The only non-OECD country in the sample is Brazil, for which the UI policy has been recently studied too (Gerard and Gonzaga 2016; Gerard and Nairitomi 2019). The approach I take and the assumptions I have made for each country are explained in detail in Web Appendix A and the calculations
are detailed in Web Appendix B.

Figure 1 plots for these 18 different countries the replacement rate at the start of the unemployment spell for a 35 year old worker in a household with an employed partner and one child. The worker earned the respective country’s average wage prior to unemployment. Large heterogeneity exists among the different countries, with replacement rates going from as low as 0.1 in the United Kingdom to as high as 0.8 in Switzerland. No country offers full replacement of pre-unemployment wages for the average worker. One can see a well-known divide in the generosity of social insurance programs, with Anglo-Saxon and Eastern European countries on the less generous side of the benefit spectrum, and Western and Northern European countries on the more generous side. Asian countries tend to have replacement rates around the average of the sample, at least for this particular scenario.

Figure 1: UI replacement rate across countries

Notes: The Figure shows the replacement rate under the same baseline scenario in all 18 countries. I consider a 35-year old worker, earning the average wage in her respectively country, fully eligible and part of a household with partner (also employed at the average wage) and one child. This scenario is changed in the following figures. Further detail is provided in the Web Appendix.

A first important determinant of the unemployment benefit level to which a worker is eligible are her earnings pre-unemployment. This is illustrated in Figure 2. I compare the replacement rate at the start of the unemployment spell for individuals who, prior to entering unemployment, earned their country’s average wage (as in figure 1) with those who earned only half of that. Note that in Sweden, the average pre-unemployment wage
of unemployed individuals is almost exactly half the average national wage (Kolsrud et al., 2018). Most countries lie below the 45-degree line in Figure 2, meaning that replacement rates tend to be higher for individuals with lower pre-unemployment wages. Typically, for those cases benefits are calculated as a percentage of pre-unemployment earnings subject to an upper cap which is binding at the average wage, but not at half of the average wage. In Spain, for example, initial unemployment benefits are set at 70% of the individual’s pre-unemployment wage, capped at the monthly amount of 1254.96 Euros, which is lower than 70% of the average wage, but not lower than 70% of half the average wage. This results in a replacement rate of less than 70% for individuals whose pre-unemployment wage was equal to the average wage. A few countries are shown precisely on the 45-degree line of the graph, meaning that the replacement rate is the same for the two considered cases of pre-unemployment wage. In those cases, the benefit cap is typically non-binding even at the average wage. This is the case of Portugal, where the 65% replacement rate is applied both at the average wage and half the average wage. For the remainder of the analysis, I consider individuals who earned half of the national average wage.

Figure 2: Cross-country UI replacement rates by pre-unemployment earnings

A second important dimension along which unemployment benefits vary is the length of the ongoing unemployment spell. Figure 3 illustrates how unemployment benefits vary over the course of the unemployment spell, comparing the replacement rates at the start of the unemployment spell with those one year into the spell. I
consider individuals who contributed to the UI system for one year, which in many countries is the minimum necessary for benefit eligibility. In general, benefits decline over the course of the spell. In fact, in many countries, UI benefits are limited in time and workers exhaust their UI benefits before one year of unemployment. In the US for example, workers tend to receive UI benefits for only six months. In several countries, however, workers are eligible to further Unemployment Assistance, typically characterized by smaller amounts, no contribution requirements and no duration limit. This is the case, for example, in Germany. At the start of the unemployment spell, a German worker with dependent children has the right to claim benefits worth 67% of pre-unemployment wage (subject to a cap which is non-binding here), but after one year such eligibility has expired. The individual is then eligible for further unemployment assistance, which comprises flat payments that amount to less than 67% of pre-unemployment wage. The Unemployment Assistance is often means-tested. For example, in the UK, unemployed workers exhaust Job Seeker’s Allowance after six months unless their household income and savings are below specific thresholds.

Figure 3: Cross-country UI replacement rates by unemployment duration

Figure 4 illustrates how contribution time and age, often interacted, impact benefits over the course of the unemployment spell. I compare the replacement rates for unemployed workers who are one year into the spell for two situations: younger workers who contributed to the UI system for only 1 year and older workers who contributed for 5 years. In certain countries, benefits one year into the spell are the same in both
situations, either because all eligible payments are exhausted (e.g., Japan and Brazil), or because the worker has moved away from regular UI benefits into further unemployment assistance which does not depend on age and contribution (e.g., Greece and Austria). On the other hand, it is often the case that older individuals with a longer contribution history can claim higher benefits in the long-run. For example, in Italy, older individuals may receive regular UI for longer than one year, while younger workers don’t. In Spain, it is the contribution history that determines the duration of regular UI benefits: an individual who contributed for five years can claim UI for longer than a year, while an individual who contributed for only one year moves to less generous unemployment assistance before the end of the first year of the unemployment spell.

Figure 4: Cross-country UI replacement rates for LT unemployed with varying years of contribution and age

To further illustrate how the differentiation of the UI policy based on individual characteristics can vary across countries, I show the UI generosity in Sweden and Switzerland side-by-side under different scenarios in Figure 5. In both countries, a worker earning half the national average wage, aged 35 with one child, is entitled to the same level of benefits at the start of the spell, 80% of pre-unemployment wages. However, different individual characteristics have different consequences for the level of benefits in both countries. For a worker with pre-unemployment wage equal to the full national average wage, the replacement rate falls to 50% in Sweden, but remains at 80% in Switzerland. In both countries, benefits decrease after one year if the worker
has only contributed for one year. However, for older workers with a longer history of UI contributions, benefits do not decrease after one year in Switzerland, but do so in Sweden. In addition, Switzerland has a child premium, with the replacement rate for a childless worker falling to 70%, while in Sweden the presence of a child does not impact the benefit amount.

Figure 5: Within country variation: comparing Switzerland and Sweden

This brief tour has highlighted some important drivers of variation in insurance coverage and how their importance differs in different countries, but this does not do full justice to the often complex myriad of rules that countries apply and how these rules differ across countries. Not only the earnings and employment history before unemployment, but also the type of job separation determines the UI generosity. In France, the benefits depend on whether the unemployment spell follows the end of fixed-term contract, an individual layoff or an economic layoff. In most countries, only involuntary quits are covered by the UI program. In Sweden, however, workers can receive unemployment benefits after voluntary quits, but only 50 days after unemployment. Waiting periods are also used for young workers who have just entered the labor market (e.g., Belgium). In addition to the unemployment durations, countries like Brazil and Portugal make the level of UI benefits dependent on the number of past unemployment spells. In the US, the potential benefit duration can be extended for workers when the unemployment rate is high. The effects of this countercyclical UI policy has been the subject of a recent debate (Farber and Valletta, 2015; Chodorow-Reich and Karabar-
In addition to extra eligibility conditions, there are other transfers than the unemployment benefits that the unemployed may specifically be eligible to. One example is financial aid for housing and utilities, which can be sizeable. In Germany, for example, a couple with one child can receive up to 562 Euros extra per month.

3 UI Design in Theory

This section revisits the Baily-Chetty characterization of optimal UI, but allowing the UI benefit level to depend on individual characteristics. In the spirit of the sufficient-statistics approach, the characterization is a function of few high-level moments that can in principle be estimated empirically, which is the topic of the next section. While I touch upon some of the issues here, I refer the interested reader to Chetty and Finkelstein (2013) and Kleven (2018) for a more in-depth discussion of the Baily-Chetty formula and the so-called ‘sufficient statistics’ approach and its challenges. The extension and notation I use builds on Kolsrud et al. (2018).

3.1 Setup

We consider an unemployment policy designed to protect workers against unemployment risk. The unemployment policy $P = \{b_x\}_{x \in X}$ conditions the UI benefit level on a vector of characteristics $x$. This could include the workers’ pre-unemployment earnings, the unemployment duration, the age at the onset of the spell, macro-economic indicators, etc. A key concern in the design of the policy is that the share of unemployed workers will depend on the generosity of the policy and thus affects the expected cost of the policy. Denoting the total labor force by $L$ and the share of workers on unemployment benefit $b_x$ by $S_x$, one can write the government’s budget surplus as

$$G(P) = [\bar{\tau} - \sum_{x \in X} S_x [b_x + \tau_x]] \times L,$$

where $\bar{\tau}$ is the average tax in the labor force and $\tau_x$ the average tax for workers with characteristics vector $x$. The fiscal cost of unemployment not only depends on the UI benefits $b_x$, but also on the foregone tax revenues $\tau_x$.

Not only the share of unemployed workers will depend on the policy. The fact that the variables used to differentiate the policy can be endogenous to the policy itself matters for the evaluation of a differentiated unemployment policy. For example, when UI benefits are more generous for higher pre-unemployment wages, this may affect workers’ incentives to get a high-wage job. Similarly, when UI benefits are less generous for long-term unemployed, this changes workers’ incentives to leave unemployment early in the spell. Even for seemingly exogenous characteristics like age used to differentiate UI benefits, workers can still manipulate the age which they start an unemployment
spell (Doornik et al., 2018; Citino et al., 2019). That is, when the potential benefit duration is longer for older workers, this affects workers’ incentives to delay the start of an unemployment spell. In general, the share of workers on $b_x$ will not only depend on the efforts of workers’ with characteristics $x$ to avoid unemployment, but also on workers’ incentives to be eligible for $b_x$. As the setup is general, the insights extend beyond the differentiation of standard UI benefits and can account for other social assistance unemployed workers can benefit from (e.g., housing benefits) or alternative policies individuals who are out of work can consider applying for (e.g., disability benefits, early retirement benefits).

I assume that any worker tries to maximize her welfare given the policy in place, affecting potentially her unemployment risk and UI benefit eligibility, and denote the resulting utility by $V_i(P)$. Social welfare associated with an unemployment policy $P$ can then be written as

$$W(P) = \int V_i(P) \, di + \lambda \left[ G(P) - G \right], \quad (2)$$

where $G$ is an exogenous revenue constraint and $\lambda$ equals the marginal cost of public funds. The policy’s central trade-off is to provide insurance against unemployment while maintaining incentives to avoid unemployment. The characterization naturally extends for general Pareto weights to account for redistributive motives.

### 3.2 Sufficient Statics Approach

The most powerful idea - but also its strongest limitation - underlying the sufficient statistics approach is to focus on local deviations from the current policy. The welfare impact of local deviations are easier to characterize as we only need to account for the externalities of individuals’ responses to the policy change, as we discuss below. This can predominantly depend on the corresponding impact on the government’s budget constraint, which workers do not internalize. As is well known, local deviations from the current policy can be considered to test the optimality of the policy in place, as no deviation from an optimal policy can increase welfare. But local deviations also allow to identify the direction in which a suboptimal policy should be changed to increase welfare. However, the important caveat is that we are restricted to local recommendations. This can be particularly restrictive when evaluating multi-dimensional policies: the welfare impact of changing $b_x$ may crucially depend on $b_{x'}$. We would need to embed the approach in a more structural framework to evaluate big policy reforms or to provide recommendations on the optimal policy. However, the workings of the structural framework will depend on model primitives that are harder to identify, leading to a trade-off between the two approaches (Chetty, 2009). The frontier is to leverage the advantages of both approaches, for example using the sufficient-statistics approach to identify and estimate the welfare impact of local changes and using the structure of
the model to gauge how this welfare impact would change as we move away from the current policy (Kolsrud et al., 2018).

**Local Deviations**  Consider now an increase in the benefit level \( b_x \), keeping all other parts of the unemployment policy fixed. The total impact on welfare depends on how much the unemployed who benefit from this increase value it in comparison to its budgetary cost. This can be expressed as

\[
\frac{\partial W(P)}{\partial b_x} = \int \frac{\partial V_i(P)}{\partial b_x} di + \lambda \frac{\partial G(P)}{\partial b_x}.
\]

(3)

As well-known now in the public finance literature, the welfare effect of a policy change depends on the agents’ behavioral responses, but only to the extent that the agents’ behavior has consequences that they did not internalize themselves. Indeed, invoking the envelope theorem, an agent’s response to a policy change will have only a second order impact on her own welfare \( V_i(P) \) (Chetty, 2006). For example, the agent may change her job search strategy in response to an increase in UI benefits, but since she was optimizing before the policy change, this response will have a negligible effect on her own welfare. Similarly, the agent may undertake action to become eligible for \( b_x \), but since she was making this trade-off optimally before the policy change, the welfare impact will be of second order.\(^1\) As a consequence, in the absence of other externalities, we only need to account for the impact of behavioral responses on the government’s budget \( G(P) \) - the fiscal externality - and the direct impact of the policy change on agents’ welfare.

**Value of UI**  The direct effect of an increase in \( b_x \) depends on the agents’ welfare gain from having the extra resources available when unemployed and eligible for \( b_x \). This gain is fully captured by the marginal utility of consumption \( v'_u(c_u) \) for each unemployed worker on \( b_x \),

\[
\int \frac{\partial V_i(P)}{\partial b_x} di = S_x \times E \left[ v'_u(c_u) \right].
\]

(4)

Clearly, the value of UI will depend on the severity and persistence of the income shock unemployed workers are exposed to and the private means and social transfers they have access to in order to protect themselves against such shock. Both will affect the extent to which workers can smooth consumption when unemployed. However, the utility gain from the extra dollar of consumption when unemployed, possible due to the UI benefit increase, is all we need to know to capture the value of UI at the margin. The overall effect thus simplifies to the share of workers who are unemployed and eligible to \( b_x \) and

\(^1\)Changes in the choice variables might be discontinuous in response to small policy changes. In principle one can allow for such discontinuous behavioral responses if they average out when integrating across heterogeneous individuals so that the social welfare function is differentiable.
their average marginal utility of consumption. Note that the income shock due to job loss may well extend beyond the unemployment spell as unemployed workers may get re-employed at lower wages and in more insecure jobs. A worker’s consumption as she goes in and out of unemployment will be interdependent. Nevertheless, as UI benefits condition on unemployment, the workers’ consumption when unemployed is sufficient to determine the marginal value of these UI benefits.

Cost of UI  Let us now turn to the budgetary impact from an increase in $b_x$. The first effect from increasing the benefit level is mechanical and again depends on the share of unemployed workers on $b_x$. The second effect is behavioral and is determined by the corresponding budgetary cost. Conveniently, there is no need to know any individual responses, since only the aggregate effect on the government’s budget matters. The budget will, however, be affected both through the increased unemployment risk in response to the more generous benefit, but also through the changed eligibility for the different benefit levels. Hence, we need to know the full vector of elasticities $\varepsilon_{S_x',b_x} = (\partial S_{x'}/\partial b_x) / (S_{x'}/b_x)$ capturing the responses of the worker shares on unemployment benefit $b_{x'}$ when changing benefit level $b_x$,

$$\frac{\partial G(P)}{\partial b_x} = -S_x - \Sigma_{x'} \frac{\partial S_{x'}}{\partial b_x} (b_{x'} + \tau) = -S_x \times \left[ 1 + \Sigma_{x'} \frac{S_{x'} (b_{x'} + \tau)}{S_x b_x} \varepsilon_{S_x',b_x} \right]. \quad (5)$$

While in principle the full set of responses determines the fiscal externality, the set of economically relevant elasticities may be smaller. Mapping out these interaction effects and estimating the ones that are relevant will be key to evaluate a differentiated UI program.

Uniform Policy  I first put together the gain and cost from increasing UI generosity for the uniform profile $b_u$, allowing us to retrieve the well-known formula in Baily (1978) and Chetty (2006). The optimal uniform policy is characterized by

$$\frac{\partial W}{\partial b_u} = 0 \iff \frac{E \left[ v'_u (c_u) \right]}{\lambda} = \frac{b_u + \tau_u}{b_u} \varepsilon_{S_u,b_u}. \quad (6)$$

The left-hand side measures the premium society is willing to pay, as expressed by the social welfare function, to transfer an extra dollar to an unemployed worker. This crucially depends on the difference in marginal utility of consumption when unemployed and when employed. The more workers can smooth their consumption when losing

---

2 In comparison, the marginal value of severance payments would be determined by the consumption at layoff, regardless of unemployment spell that follows. Or, the marginal value of re-employment bonuses would be determined by the consumption on re-employment.

3 The elasticities are weighted by the relative share of the budget spent on different parts of the unemployment policy. The budgetary spillover effects of a change in $b_t$ on other parts of the policy is less relevant the less generous these other parts are. There is, however, a correction for the tax rate because more time spent unemployed also reduces the taxes received from employment.
their job, the less valuable UI is. The right-hand side measures the premium society ought to pay when transferring an extra dollar to the unemployed due to the increase in unemployment. The higher the elasticity of the unemployment rate, the more costly UI is. At the optimum, the marginal value and cost of UI should be equalized.

3.3 Differentiated Unemployment Policy

We first generalize the characterization of the optimal uniform policy to the optimal differentiated policy when workers have no control over the eligibility criteria. That is, \( \varepsilon_{S',x} = 0 \) for \( x' \neq x \). For each part of the policy, we need

\[
\frac{\partial W}{\partial b_x} = 0 \iff \frac{E[v'_u(c_u)|x] - \lambda}{\lambda} = \frac{b_x + \tau_x}{b_x} \varepsilon_{S_x,b_x}.
\]  

(7)

The consumption smoothing gains and the unemployment response among the group of workers eligible to \( b_x \) are sufficient to evaluate whether it is optimally set. Put differently, for two different groups of workers, any differentiation of the UI policy is only justified if either the consumption smoothing gains or the unemployment response is different. The test for whether UI should be more generous for one group relative to the other is simply whether the consumption smoothing gains relative to the unemployment cost are higher.

The difference in consumption smoothing gains and unemployment responses across different groups of unemployed workers remains key when the unemployment policy is differentiated based on endogenous conditions. To characterize the optimal differentiated policy, we do, however, need to account for the impact of policy changes across different parts of the differentiated policy,

\[
\frac{\partial W}{\partial b_x} = 0 \iff \frac{E[v'_u(c_u)|x] - \lambda}{\lambda} = \frac{b_x + \tau_x}{b_x} \varepsilon_{S_x,b_x} + \sum_{x' \neq x} S_{x'} (b_{x'} + \tau_{x'}) S_{x'} \varepsilon_{S_{x'},b_x}.
\]

(8)

I provide two examples to shed light on the potential role of these cross-elasticities.

Consider first a two-part benefit policy differentiating the UI benefit level for two groups of workers: e.g., low vs. high-income workers, \( X = \{ y < \bar{y}, y \geq \bar{y} \} \). The share of workers on \( b < \bar{y} \) depends on the share of low-income workers and the unemployment rate among low-income workers, \( S_{<\bar{y}} = F_{<\bar{y}} U_{<\bar{y}} \). A change in UI for low-income workers has the following budgetary effect,

\[
\frac{\partial G(P)}{\partial b_{<\bar{y}}} = -S_{<\bar{y}} - \frac{\partial S_{<\bar{y}}}{\partial b_{<\bar{y}}} (b_{<\bar{y}} + \tau_{<\bar{y}}) - \frac{\partial S_{\geq\bar{y}}}{\partial b_{<\bar{y}}} (b_{\geq\bar{y}} + \tau_{\geq\bar{y}}) \approx -S_{<\bar{y}} - F_{<\bar{y}} \frac{\partial U_{<\bar{y}}}{\partial b_{<\bar{y}}} (b_{<\bar{y}} + \tau_{<\bar{y}}) - \frac{\partial F_{<\bar{y}}}{\partial b_x} U_{\bar{y}} (b_{\geq\bar{y}} - b_{<\bar{y}}),
\]

(9)

where the approximation relies on the effect of the low-income benefit level on the high-income unemployment rate to be small (\( \partial U_{\geq\bar{y}}/\partial b_{<\bar{y}} \approx 0 \)). We can obtain a similar
expression for $b_{\geq \bar{y}}$. The endogeneity of the condition used to differentiate UI requires accounting for the share of workers who may lower their income to become eligible for the more generous unemployment benefit level for low-income workers. The fiscal externality depends on the unemployment rate among these workers at the margin, which I denote by $U_{\bar{y}}$, and the difference in UI generosity on the two parts of the policy. Importantly, starting from a uniform policy $b_{\geq \bar{y}} = b_{< \bar{y}}$, this fiscal externality becomes of second order, and so even with endogenous conditions, it is only the unemployment response of the directly affected group that matters when considering to differentiate the UI generosity of that group. By the same token, the more differentiated the UI policy, the more important the eligibility responses are when considering to further differentiate the policy.

Consider now a two-part benefit policy differentiating the UI benefit level for two parts of the unemployment spell: e.g., the first six months of unemployment and thereafter, $X = \{d < \bar{d}, d \geq \bar{d}\}$. This is the setting analyzed in Kolsrud et al. (2018). The share of unemployed workers on the different parts of the UI policy are now directly related as the share of long-term unemployed depends on the share of short-term unemployed and their exit rate out of unemployment. More formally, we have $S_{d+1} = S_d (1 - h_d)$, where $h_d$ is the hazard rate at unemployment duration $d$. The corresponding cross-elasticities have been at the heart of the theoretical models used to study the optimal timing of UI benefits. On the one hand, generous unemployment benefits at the start of the spell ($b_{< \bar{d}}$) will discourage workers from finding employment and affect the share of workers moving on to the second part of the unemployment policy. On the other hand, generous unemployment benefits later in the spell ($b_{\geq \bar{d}}$) already reduce the incentives early in the spell for job seekers who are forward-looking and contemplate the risk of being long-term unemployed. In addition to the cross-elasticities, there are of course the direct effects: how responsive are short-term vs. long-term unemployed workers to changes in unemployment benefits? And how do these responses compare to consumption smoothing gains from the short-term and the long-term unemployed? The power of the proposed approach is to capture all the forces together and turn the question into an empirical one. In fact, when starting from a flat benefit profile $b_u$, the characterization simplifies to

$$\frac{\partial W}{\partial b_{< \bar{d}}} = 0 \iff \frac{E \left[ v' \left( c_u \right) \mid d < \bar{d} \right] - \lambda}{\lambda} = \frac{S_{< \bar{d}} + S_{\geq \bar{d}}}{S_{< \bar{d}}} \frac{b_u + \tau_u}{b_u} \frac{c_{\left[S_{< \bar{d}} + S_{\geq \bar{d}}\right]}}{b_{< \bar{d}}}. \quad (11)$$

The benefit level at a given unemployment duration should be lower, when the consumption smoothing gains are lower and when the overall unemployment response to changes in that benefit level is higher, accounting for the share of workers that are still unemployed at that duration.
4 From Theory to Practice

The key advantage of the identified trade-off between incentives and insurance a la Baily-Chetty is its potential to be evaluated empirically. The identified moments have clear empirical counterparts, which make the link from theory to recommendations very transparent. However, while the characterization is general and robust to richer models, the empirical implementation naturally requires assumptions. These assumptions can be strong, so it is first of all important to be transparent on them. Moreover, it is valuable to use complementary approaches and data to try to relax the assumptions or consider policy recommendations that do not depend on these assumptions. I discuss this briefly in the context of differentiated unemployment policies.

4.1 Consumption Smoothing

The value of higher unemployment benefits to a group of workers depends on the marginal utility of consumption for that group of workers when unemployed. It is standard in the literature to ignore the redistributive value of unemployment insurance, which would affect the wedge between $E[v'_u(c_u)|x]$ and the marginal cost of public funds $\lambda$. The focus is instead on the insurance value, comparing the marginal utility of consumption when unemployed vs. employed for a given worker. At the margin, the insurance value of transferring an additional dollar from employment to unemployment, depends on the drop in consumption that an unemployed worker is exposed to and how much she cares about the variation in consumption. The standard consumption-based approach in the literature is to estimate the drop in consumption when unemployed and then scale the drop in consumption with a risk preference parameter. Indeed, assuming CRRA preferences with $\gamma = -cv''(c)/v'(c)$, the relative difference in marginal utilities simplifies to:

$$\frac{v'(c_u) - v'(c_e)}{v'(c_e)} \approx \frac{c_e - c_u}{c_e}. $$

Implementing the consumption-based approach is remarkably easy as it only requires linking data on consumption expenditures to data on unemployment status. In recent years more registry-based data on consumption expenditures have become available, which can be linked to employment registers as well (Kolsrud et al., 2018, Gerard and Naritomi, 2019). A major advantage for the implementation is that in principle there is no need to know the means that workers use to smooth their consumption. This is, however, conditional on knowing the preference parameter with which the consumption drop should be scaled. The relevant preference parameter would depend on whether workers’ preferences over consumption are state-dependent, for example through complementarities with leisure, but also on whether observable consumption expenditures are state-dependent and affect welfare-relevant consumption differently when employed and unemployed. This is a specific concern in the context of unem-
ployment because of differences in work- or job-search related expenditures or because unemployed workers can substitute towards home production and shop at lower prices. More generally, the relevant preference parameter will depend on the type of consumption expenditures that unemployed job seekers can change. For example, being able to lower expenditures on durable goods will affect the marginal utility of consumption less (Browning and Leth-Petersen 2003), while being forced to only lower non-committed expenditures will affect the marginal utility of consumption more (Chetty and Szeidl 2007).

The challenge of translating consumption wedges into marginal utility wedges does not disappear when evaluating the optimal differentiation of a policy, but the nature of the challenge changes. While differences in consumption drops across workers with different characteristics may seem indicative, one needs to know how preferences change as well. For example, to evaluate how high unemployment benefits should be for workers with low pre-unemployment earnings, we should consider the consumption drop in this group, but we would also need to know their respective preferences. In fact, differences in consumption may be more than offset by differences in preferences, as workers with different preferences will invest differentially in consumption smoothing (Chetty 2006; Landais and Spinnewijn 2018). However, to evaluate the relative generosity of two parts of the unemployment policy, we no longer need to know the level of the preference scalars $\gamma_x, \gamma'_x$, but it is sufficient to know the relative preferences $\gamma_x / \gamma'_x$. In particular, when we know that preferences are comparable $\gamma_x \approx \gamma'_x$, it becomes sufficient to know the relative drop in consumption. Kolsrud et al. (2018) find that the consumption drop is more pronounced for the long-term unemployed than for the short-term unemployed, consistent with the fact that workers run out of assets as they remain unemployed for longer. This indicates that the consumption smoothing gains are larger for the long-term unemployed, unless the workers who select into long-term unemployment are less averse to the reduced consumption.

Converting observable responses or wedges into comparable objects for welfare analysis will always be challenging. The extra challenge in the context of UI is that coverage is mandated and workers do not directly reveal their valuation through their coverage choice. A number of recent papers in the UI literature further tackle this challenge proposing alternative approaches. In the spirit of the consumption-based approach, one can look at wedges in resources used when employed and unemployed, for example changes in spousal labor supply (Fadlon and Nielsen 2018), or consider responses to changes in the anticipated unemployment risk (Hendren 2017). These responses still need to be scaled by a preference parameter. A way to circumvent the scaling is to use differences in behavioral responses to different sources of income variation. Chetty

---

4 See Chetty and Finkelstein (2013) and Landais and Spinnewijn (2018) for further discussion of the challenges for the consumption-based approach.

5 An exception are the UI policies in Scandinavian countries where workers are offered to choice between basic and more comprehensive coverage (Landais et al. 2017).
 show how the differential response in unemployment risk to changes in unemployment benefits relative to other income changes can capture consumption smoothing gains. Landais and Spinnewijn (2018) instead show how to use the differential marginal propensity to consume when unemployed and employed to identify the revealed cost of smoothing consumption so as to bound the value of consumption smoothing. Overall, the papers using the alternative approaches suggest that the insurance value of UI is higher than what the consumption-based approach suggests (using ‘standard’ values of risk aversion). The implied mark-ups that workers are willing to pay to transfer consumption from employment to unemployment are closer to 100% and sometimes above, rather than around 10%−50% as suggested by the consumption-based approach. These differences are sizeable, but very little is known on how the estimated consumption gains differ for different groups.

## 4.2 Unemployment Response

In comparison with the literature studying the insurance value, there is a vast literature estimating the labor supply effects of social insurance (Krueger and Meyer, 2002), and the unemployment responses to changes in UI benefits in particular (Schmieder et al., 2016). The predominant focus in the literature is on the unemployed themselves and how unemployment duration changes when varying the benefit level or the potential benefit duration. To estimate the impact of UI generosity, it is common to use exactly the variation that comes from the differentiation in unemployment benefits, for example by pre-unemployment earnings, by contribution years, or by age at layoff. The differentiated schedules indeed include jumps or kinks for example at earnings thresholds or age cut-offs, providing plausibly exogenous variation. The variation, however, does not allow to evaluate the differentiated schedule itself.

Let me illustrate this for a two-part benefit policy differentiating the UI benefit level for low- vs. high-income workers, $X = \{y < \bar{y}, y \geq \bar{y}\}$. Assume there is a jump in the benefit level from $b_{<\bar{y}}$ to $b_{\geq\bar{y}}$ at income threshold $y = \bar{y}$. A standard approach in the literature (Card et al. 2007; Schmieder et al. 2012) is to use a regression-discontinuity design linking the difference in unemployment outcomes just above and below the income threshold, $\lim_{y \to \bar{y}^+} U_y - \lim_{y \to \bar{y}^-} U_y$, to the jump in the UI benefits at the threshold, $b_{\geq\bar{y}} - b_{<\bar{y}}$. To interpret this estimate as the causal effect of benefits on unemployment for workers at the income threshold, $\partial U_{y=\bar{y}}$, we need that workers around the threshold are similar, except for the UI benefit they are eligible to. To gauge this identifying assumption, a standard check is whether workers are comparable on

Landais and Spinnewijn (2018) also leverage the UI choices in Sweden to obtain revealed-preference estimates of the value of insurance, and implement the various approaches in the same context on the same sample of workers.

Card et al. (2015) and Kolsrud et al. (2018) use the kink in the benefit schedule as a function of pre-unemployment earnings due to a cap on UI benefits. Card et al. (2007) and Schmieder et al. (2012) use jumps in the benefit schedule at age and contribution year cut-offs.
observables and whether workers do not change their income in response to the benefit jump (McCrary, 2008).

The object of the regression-discontinuity design is quite different from what is needed to evaluate the differentiated policy. First, the response in earnings to benefit differences below or above an earnings threshold is not a threat to identification, but one of the response of interest as the corresponding spill-over effects ($\frac{\partial F_{< \bar{y}}}{\partial b_{< \bar{y}}}$ and $\frac{\partial F_{\geq \bar{y}}}{\partial b_{\geq \bar{y}}}$) need to be accounted for. If there is a non-zero response, it is the average unemployment rate for the workers who change their income in response to the benefit change that is relevant relevant, as these responding workers increase the UI expenditures by $b_{\geq \bar{y}} - b_{< \bar{y}}$ when they become unemployed.\(^8\) If there is no response, the evaluation of the differentiated policy simplifies to the standard Baily-Chetty expression. Second, to evaluate the differentiated policy, the value of knowing workers behavior at the threshold, i.e., $\frac{\partial U_{y=\bar{y}}}{\partial b_{y=\bar{y}}}$, is fairly limited. What we need to know instead is how the unemployment response to benefit differs for different income groups ($\frac{\partial U_{< \bar{y}}}{\partial b_{< \bar{y}}}$ vs. $\frac{\partial U_{\geq \bar{y}}}{\partial b_{\geq \bar{y}}}$). This requires independent variation in benefits given to low-income workers $db_{< \bar{y}}$ and high-income workers $db_{\geq \bar{y}}$ respectively. The jump in the benefit schedule does not provide this variation.\(^9\)

In general, variation in the differentiated benefit schedule is hard to come by. A notable exception is Kolsrud et al. (2018) who exploit variation in UI benefits, both early in the spell and late in the spell, to evaluate the time profile of UI benefits. The variation is driven by a cap on unemployment benefits, which depends on the spell duration and changes over time, providing independent variation in $db_{< \bar{d}}$ and $db_{\geq \bar{d}}$. This allows estimating all relevant cross-elasticities, $\varepsilon_{S_x, b_x}$ for $x, x' \in X = \{d < \bar{d}, d \geq \bar{d}\}$, to evaluate the optimal differentiation between short-term and long-term unemployed.

The large empirical literature on the unemployment effects of UI benefits has provided a wide range of elasticity estimates, ranging by and large from .5 to 1.5. Some of the differences in estimates can be simply reconciled by the potential duration over which the benefits are changes. The largest estimates are found when the potential benefit duration is long and the overall benefit level is increased (Card et al., 2015; Kolsrud et al., 2018). In the US, where benefits are exhausted after six months of unemployment, the responses to variation in unemployment benefit levels tend to be larger (Meyer, 1988; Landais, 2015) than the responses to changes in the potential benefit duration (Rothstein, 2011; Farber and Valletta, 2015). This is also consistent with the findings in Kolsrud et al. (2018) who estimate the incentive costs to be larger for UI benefits paid to the short-term unemployed compared those paid to long-term unemployed. This difference holds despite the forward-looking response by short-term unemployed to changes in UI benefits when long-term unemployed.

\(^8\) An important concern in the presence of selection in an RDD is selection on treatment (Gerard et al., 2016), but this is only of second order importance when evaluating the budgetary effect of small changes in $b_x$.

\(^9\) The identification issues here also relate to the distinction between the LATE, which the variation in UI policies allows us to estimate, and the heterogeneity in the Marginal Treatment Effects, which we would need to evaluate the heterogeneous responses.
**Differentiated Unemployment Policy** Putting the estimates of the insurance value and incentive costs together, we should be able to conclude whether UI benefits are too high or too low. But reaching consensus can be challenging. Compared to the traditional consumption-based estimates of the insurance gains, even the low-range estimates of the incentive costs imply that UI would be too generous overall. However, the more recent estimates on the insurance gains indicate that even for high incentives costs, a further increase in generosity may be desirable.

Linking this back to the dimensions of policy differentiation discussed in Section 2, it is also crucial to get a better understanding of the heterogeneity in consumption smoothing gains and unemployment responses. We currently have very little evidence on the corresponding heterogeneity by pre-unemployment earnings, by pre-unemployment history, by age and by available means, all common dimensions across which UI generosity differs. Moreover, we would need to systematically compare the heterogeneity in impacts to the eligibility responses to the differentiation in the policy itself.

The evidence in Kolsrud et al. (2018) allows to evaluate the optimal differentiation by unemployment duration and illustrates that the challenges for drawing conclusions on the optimal differentiation are different from the challenges for pinning down the optimal level. In particular, as the incentive costs are larger for UI benefits paid to the short-term unemployed, while the consumption drops are more pronounced for the long-term unemployed, the robust conclusion is that benefits should be higher for the long-term unemployed, unless the long-term unemployed have very different preferences (or are assigned different social welfare weights). This is still a local recommendation, but does not rely on whether on average UI benefits are too high or too low. Hopefully future research can identify other significant dimensions of heterogeneity and shed further light on the optimal differentiation of UI.

5 Discussion

A central challenge in the design of UI is to protect workers against unemployment risk while providing incentives to reduce unemployment risk. In practice, countries provide different protection to workers with different characteristics or employment histories. This differentiation of UI could be evaluated according to the same trade-off, where the UI generosity should be higher for workers who value insurance more or for whom the incentive cost is lower. The benefit from differentiating the generosity will be higher when there is more heterogeneity in valuation relative to cost, assuming that workers have limited ability to change their eligibility status. More empirical work is needed

Landais and Spinnewijn (2018) document some observable heterogeneity in the consumption drops and the value of UI as revealed by workers’ UI choices. The correlation in the heterogeneity in both measures is often insignificant or even negative, indicating the potential importance of heterogeneity in preferences or the confounding role played by frictions underlying choices.
to evaluate the current differentiation in UI benefits - in particular the differentiation by unemployment duration and by pre-unemployment earnings. But we should also go a step further to uncover key dimensions of heterogeneity in welfare impacts that are currently ignored in UI policies. For example, unemployment benefits may be better targeted when conditioning on the unemployment history rather than just on the length of the ongoing unemployment spell.

While this discussion has focused on standard UI, the evaluation could be extended to other unemployment policies. As mentioned before, severance pay - only conditioning on layoff - is relatively more common in developing countries (Gerard and Naritomi, 2019). One can compare the insurance loss with the incentive gain from such transfers - unconditional on being unemployed. Private UI savings accounts are a related policy, further reducing incentive costs, but still providing access to liquidity upon layoff. In general, the need for protection against the consequences of job loss may extend beyond the initial unemployment spell, as evidenced by the long-lasting impact of job loss on earnings (Jacobson et al., 1993) and consumption (Landais and Spinnewijn, 2018). This may call for more targeted policy instruments like wage insurance.

In principle, heterogeneity in valuations could also be accommodated by offering workers the choice over how much UI coverage to get. However, realizing these gains is not obvious when workers also differ in costs due to adverse selection or when their choices are subject to behavioral frictions (Landais et al., 2017; Landais and Spinnewijn, 2018). With the exception of some of the Scandinavian countries, in all countries the UI coverage is pre-determined, leaving no choice to the worker.

Just like most of the UI literature, I have restricted my analysis of the gains and costs of UI to the insurance value and the incentive cost respectively. In practice, unemployment insurance may be used for fiscal stabilization (Kekre, 2017) and to stimulate consumption (McKay and Reis, 2016). The general equilibrium effects may be different from the partial unemployment responses, which affects the characterization of the optimal trade-off (Landais et al., 2018a; b). For evaluating the optimal differentiation of UI, the question becomes whether these general equilibrium effects are different for the unemployment benefits paid to different groups. My discussion has also overlooked the role of behavioral frictions, although they are shown to be important for both workers’ job search decisions and their consumption smoothing. Examples are reference-dependence (DellaVigna et al., 2017) impatience and/or excess sensitivity (Gerard and Naritomi, 2019; Ganong and Noel, 2017), biased beliefs (Spinnewijn, 2015; Mueller et al., 2018). When evaluating different types of unemployment policies, it is crucial to understand how they affect workers’ exposure to these biases.

When comparing how desirable a policy is for different groups of workers, only the relative incidence of these biases will be important. So even when accounting for these additional effects, a robust evaluation of the differential protection embedded in the differentiated UI policies seems valuable and feasible.
References


