Some Observations on "The Limited Macroeconomic Effects of Unemployment Benefit Extensions" and Related Writings by Gabriel Chodorow-Reich and Loukas Karabarbounis

Marcus Hagedorn[*]
Iourii Manovskii[†]
Kurt Mitman[‡]

August 1, 2016

Abstract

In Hagedorn et al. (2016a) we reviewed Chodorow-Reich and Karabarbounis (2016a) among other recent contributions to the literature studying the aggregate implications of unemployment benefit extensions. We were also invited to discuss their paper at the summer 2016 NBER EF&G meeting with the slides of our discussion available as Hagedorn et al. (2016b). We found the identification argument in CRK (2016a) to be incorrect because the exogeneity of the estimator they proposed and implemented does not follow from their fundamental identifying assumption of exogeneity of the measurement error in unemployment.

The authors responded to our assessment in Chodorow-Reich and Karabarbounis (2016b,c) and proposed a new identification argument to justify their measurement strategy. In this paper we explain that their new identification argument does not overcome the endogeneity problems we previously documented. To avoid being repetitive, we focus only on the issues relevant for understanding CRK's two identification strategies and their replies. We refer the interested reader to HMM (2016a,b) for additional thoughts on CRK's analysis. A complete replication package of all the data, programs, and log files reported in this paper is available to download on our websites.

[*]University of Oslo, Department of Economics, Box 1095 Blindern, 0317 Oslo, Norway. Email: marcus.hagedorn07@gmail.com.

[†]University of Pennsylvania, Department of Economics, 160 McNeil Building, 3718 Locust Walk, Philadelphia PA 19104. Email: manovski@econ.upenn.edu.

[‡]Institute for International Economic Studies, Stockholm University, SE-106 91 Stockholm, Sweden. Email: kurt.mitman@iies.su.se.

1 Outline

The foundation of any empirical paper is the identification argument. In particular, to establish identification of economic effects of unemployment benefit extensions in the data, a researcher must prove that the variation in benefit duration exploited in the study is *exogenous* to the outcome variable of interest. The best proof is theoretical, but it may also be based on verifying that the exogeneity restriction is satisfied in the data.

CRK (2016a) proposed such an identification strategy based on the following logic. Unemployment is measured in real time with an error. This error is exogenous to true economic conditions. As benefit duration is based on real-time unemployment, errors in measured unemployment give rise to UI benefit errors that are orthogonal to true economic conditions. As UI errors identified by CRK (2016a) were serially correlated, they proposed to study impulse responses to *innovations* in UI benefit errors. If UI errors are exogenous, innovations to UI errors can be expected to be exogenous as well. In short, exogenous measurement error in unemployment induces exogenous errors in benefits, and the exogeneity of benefit errors is preserved when constructing UI benefit error innovations. By studying the response of labor market variables to exogenous UI benefit error innovations, one can infer the effects of UI benefit extensions.

In HMM (2016a), we have shown that this logic is flawed because even if measurement error in unemployment is exogenous, the induced errors in UI benefits are not. We provided a formal argument for this and verified that it holds in the data. CRK (2016b,c) have not contested these findings. The endogeneity of UI errors breaks the logic of CRK (2016a)'s identification argument because the exogeneity of UI error innovations does not follow from the exogeneity of measurement errors in unemployment.

CRK (2016b) have responded with a different identification argument. The argument no longer involves the exogeneity of UI benefit errors or even the exogeneity of measurement error in unemployment. Instead, they argue that innovations to endogenous UI benefit errors can be exogenous by construction. CRK do not formalize this identification claim in their writings and explain it only by means of an example in CRK (2016b). The example shows that innovations can be *defined* to be orthogonal to economic conditions. This clearly does not achieve identification, however, because the impact of such innovations on, say, unemployment does not reveal the causal impact of benefit extensions. Instead, it only reveals that the effect of innovations on unemployment was set to zero by construction. If this example correctly reflects CRK's identification strategy, it invalidates their entire analysis.

In HMM (2016a) we already noted that one can expect the innovations to an endogenous variable themselves to be endogenous, and provided some evidence on the endogeneity of CRK's innovations. CRK (2016b,c) have not acknowledged this evidence. In what follows, we explain in more detail that CRK's claim that innovations to an endogenous variable are exogenous does not follow from standard theoretical arguments. Given that CRK provide us with no alternative theoretical guidance for why benefit error innovations can be expected to be exogenous, we proceed to investigate the validity of their identification argument empirically. We show that the assertion that UI error innovations are exogenous is clearly inconsistent with the data, invalidating the identification arguments in both CRK (2016a) and CRK (2016b). Specifically, we show that

- Theory:
  - **Innovations to an endogenous regressor are typically endogenous** (Sec. 4.1)
  - In the identification example in CRK (2016b), **innovations are constructed to be orthogonal to unemployment**. This implies that the main result of the paper that innovations have no impact on unemployment does not reflect causal inference but is achieved by construction. Deviations from this construction imply that innovations are likely endogenous. (Section 3.1)
  - Replacing the regressor with **innovations typically introduces additional biases**, especially, when the LHS variable is used to construct innovations. Timing assumptions do not provide identification in a world that is not i.i.d. (Section 4.2)
- Evidence:
  - In Section 5.1 we show that **future innovations strongly predict current unemployment**. This means that CRK's analysis either does not overcome the problem of benefit endogeneity with respect to unemployment or that economic agents can foresee the "innovations" and adjust their behavior prior to "innovations" occurring.
  - In Sections 5.2 and 5.3 we provide evidence in CRK's data that benefit error **innovations are indeed endogenous** regressors, i.e. they **inherit the endogeneity problem of benefit errors**.
  - The analysis up to this point establishes that constructing innovations to an endogenous regressor does not overcome the endogeneity problem either in theory or in practice, refuting the novel identification claim in CRK (2016b). Anticipating this, in HMM (2016b) we proposed to dispense with constructing benefit errors and benefit error innovations and to directly use measurement error in unemployment as

an instrument for benefits. This relies, however, on **CRK (2016a)'s fundamental assumption** — the difference between real-time and revised unemployment is measurement error — being correct. We test this assumption in Section 5.4 and show that it **fails**. Contrary to the claims in CRK (2016c), it is for this reason that the IV strategy produces nonsensical results, as we show in Section 5.5. Using the same metric, we show that the results in CRK (2016a,b) are equally nonsensical and this happens for the same reason. We see no modification of CRK's methodology based on errors that can overcome the fundamental problem that the errors it relies on are endogenous.

2 Overview of CRK's Measurement Approach

In US states unemployment benefit duration is extended when a deterministic function of past unemployment rates rises above a pre-determined threshold. Extensions are eliminated when the same function falls below the threshold. This mechanical relationship between past unemployment rates and current benefit duration makes it difficult to estimate the effects of unemployment benefit extensions. For example, consider regressing unemployment rate in state s at time t, $\tilde{u}_{s,t}$, on benefit duration, $T_{s,t}$ as well as state and time fixed effects:

$$\tilde{u}_{s,t} = \beta T_{s,t} + \delta_s + \delta_t + \nu_{s,t}. \tag{1}$$

Although current benefit duration is a function of only *past* unemployment rates, the shocks driving state unemployment are persistent, inducing a correlation between $\nu_{s,t}$ and $T_{s,t}$. This biases the estimated β parameter.

CRK (2016a) propose a strategy to overcome this endogeneity problem. Denote by $u_{s,t}$ the unemployment rate, as measured in real time, i.e., at time t. The duration of benefits $T_{s,t}$ depends mechanically (and discontinuously) on this measurement. The data on unemployment rates are revised over time as more data become available or measurement procedures and concepts are refined. Let $\tilde{u}_{s,t}$ be the measure of the unemployment rate as revised at some future date. Denote by $\tilde{T}_{s,t}$ the hypothetical duration of benefits which would have been implemented at date t had the revised data been available in real time (given the same mechanical rule – extended benefit triggers – mapping measured unemployment rate to benefit duration).

CRK (2016a) propose to interpret $\tilde{u}_{s,t}$ as a true measure of unemployment, while the real time measure $u_{s,t}$ is the truth plus measurement error, $\hat{u}_{s,t}$:

$$u_{s,t} = \tilde{u}_{s,t} + \hat{u}_{s,t}. \tag{2}$$

In HMM (2016b) we proposed that if this fundamental identifying assumption in CRK (2016a) is satisfied, it immediately implies that the bias in regression (1) can be overcome if $\hat{u}_{s,t}$ is used to instrument benefit duration $T_{s,t}$. This is indeed a perfect instrument: measurement errors in real-time unemployment, $\hat{u}_{s,t}$, are correlated with $T_{s,t}$ but independent of $\tilde{u}_{s,t}$.

Introducing Benefit Errors. Instead of working directly with real-time unemployment benefits $T_{s,t}$, CRK proceed to construct errors in benefits. As the revised unemployment data differ from the real time unemployment data, the actual benefit duration $T_{s,t}$ differs from the hypothetical duration $\tilde{T}_{s,t}$ by an error $\hat{T}_{s,t}$ due to the measurement error in the real time unemployment rate:

$$T_{s,t} = \tilde{T}_{s,t} + \hat{T}_{s,t}. \tag{3}$$

In HMM (2016a) we have shown that $\hat{T}_{s,t}$ is *not* an exogenous measurement error. In particular, if changes in revised unemployment cross an extension threshold and trigger a change in benefits while real-time unemployment measure does not, the benefit error $\hat{T}_{s,t}$ mechanically changes in the opposite direction from $\tilde{u}_{s,t}$. This implies a negative bias in β estimated from the following regression:

$$\tilde{u}_{s,t} = \beta\hat{T}_{s,t} + \delta_s + \delta_t + \nu_{s,t}. \tag{4}$$

Introducing Benefit Error Innovations. Instead of working directly with benefit errors $\hat{T}_{s,t}$, CRK proceed to construct "innovations" in benefit errors:

$$\epsilon_{s,t} = \hat{T}_{s,t} - \mathbb{E}_{t-1}\hat{T}_{s,t}, \tag{5}$$

and use them to assess the effects of benefit extensions by estimating, for $k \geq 0$:

$$\tilde{u}_{s,t} = \beta^k\epsilon_{s,t-k} + \delta_s + \delta_t + \nu^k_{s,t}. \tag{6}$$

CRK (2016a) do not recognize the endogeneity of $\hat{T}_{s,t}$ and motivate the use of innovations only by the desire to construct the impulse response to a serially uncorrelated innovation in the benefit error. CRK (2016b) make a different argument. Acknowledging the endogeneity of $\hat{T}_{s,t}$ pointed out by HMM (2016a), they claim that replacing an endogenous variable in the regression with its innovations actually overcomes the endogeneity problem.

Note that the consequences of replacing the benefit errors with innovations in the main regression in CRK are substantial. The large and highly statistically significant estimate of $\hat{\beta} = -0.117$ using Eq. (4) turns into an insignificant estimate of $\hat{\beta} = -0.009$ using Eq. (6). CRK (2016b) show that this even happens when they perform a placebo analysis. Below we discuss three possibilities for why this happens: (1) innovations were constructed to mechanically deliver a coefficient of zero, (2) the true effect is indeed zero and constructing innovations reveals it (but using benefit errors does not), and (3) the output of the complex procedure for constructing innovations is to a large extent noise.

3 CRK (2016b)'s Econometric Breakthrough

If CRK (2016b)'s claim that constructing innovations to an endogenous regressor overcomes the endogeneity problem is correct, then it represents an econometric breakthrough. To see this, note that a classic problem in applied work is how to estimate β in a regression

$$Y_t = \beta X_t + \nu_t,$$

where X_t is an endogenous regressor, i.e. it is correlated with ν_t. Countless researchers tackled this problem and many came up with ingenious ideas how to obtain unbiased estimates: Heckman's selection model, trying to find natural experiments, looking for strong instruments, setting up field experiments, estimating sophisticated structural models... CRK argue that there is a much simpler solution—simply replace the endogenous regressor X_t with innovations to X_t:

$$\epsilon_t = X_t - E_{t-1}X_t.$$

Already in HMM (2016a) we argued that this solution is unlikely to work because innovations in general inherit the endogeneity properties of the original regressor. CRK (2016b) express surprise and ask for a more formal argument. We will provide it next. Before doing so, however, it is instructive to understand the example that CRK (2016b) use to justify their claim in lieu of any formal proof.

3.1 Proof by Example in CRK (2016b)

The only revealing explanation of CRK's vision for how constructing innovations overcomes the endogeneity problem and why the assertion in HMM (2016a) may not apply is presented in CRK (2016b) (our emphasis added)

> We defined the innovation [$\epsilon_{s,t}$] as the unforecastable component of $\hat{T}_{s,t}$ using information available as of period $t-1$. If $\epsilon_{s,t}$ is the $t-1$ forecast error of $\hat{T}_{s,t}$, then it must be uncorrelated with information available at time $t-1$. In the standard DMP model used in CRK and in a number of papers by HMM, the unemployment rate for period t is already known at time $t-1$. Therefore, even though the correlation between the UI error $\hat{T}_{s,t}$ and the unemployment rate $u^{revised}_{s,t}$ could be negative in

> the DMP model, the **correlation between the innovation $\epsilon_{s,t}$ and $u_{s,t}^{revised}$ is zero**.
>
> [This example] shows that the discussion in HMM of $\hat{T}_{s,t}$ cannot invalidate the CRK approach because, even if one constructs examples where the UI error $\hat{T}_{s,t}$ is mechanically correlated with the revised unemployment rate, the innovation in the UI error $\epsilon_{s,t}$ can still be uncorrelated with the revised unemployment rate **by construction**. Since we use $\epsilon_{s,t}$ and not $\hat{T}_{s,t}$ in our regressions, the HMM example is not relevant in evaluating the CRK methodology.

This is insightful. CRK define innovation conditional on information available at date $t-1$. As unemployment is predetermined one period in advance, this implies that *by construction* the correlation between CRK's innovation $\epsilon_{s,t}$ and $u_{s,t}^{revised}$ is zero. Recall that the ultimate objective of CRK is to measure the effects of benefits on unemployment through a regression of $u_{s,t}^{revised}$ on $\epsilon_{s,t}$. Estimating such a regression, they will find a coefficient of zero, and conclude that benefit extensions have limited macroeconomic effects. Regardless of the true effect of benefits on unemployment! In other words, CRK assign a causal interpretation to the coefficient that they set to zero by assumption.

Note that while the example suggests that the contemporaneous effect of benefits on unemployment is set to zero by construction, it leaves open the question of how CRK deal with the endogeneity bias at various lags in the impulse response.

This example highlights the two extremes available to a researcher. By directly or indirectly conditioning on all information available at time $t-1$, the innovations are orthogonal to economic conditions by construction and the estimated zero effect of innovation on unemployment has no causal meaning. On the other end of the spectrum, one may not condition on $t-1$ information at all, work directly with endogenous UI benefit errors and obtain results that are severely negatively biased. The discussion in Footnote 3 of CRK (2016b) appears to suggests that they try to do something in between of these two extremes by not fully conditioning innovations on variables known at date $t-1$ (although the equations in CRK (2016a) state otherwise).[1] Thus, while they do lower the correlation between innovation and economic conditions, they do not drive it all the way to zero. As a result such innovations remain endogenous regressors leading to a negative bias in the estimated coefficient. The relationship between this estimated coefficient and the true effect of benefits on unemployment is unknown.

[1] It might also be helpful if the authors made public their code for constructing innovations. We appreciate that they made several key data series available online. However, the paper rests entirely on one step of constructing innovations and the authors refused to share this one (and only) piece of code that we asked for.

4 Some Theoretical Questions about the Breakthrough

4.1 A Surprisingly Simple Counterexample?

Let us consider the performance of CRK's methodology in the most basic setting. For simplicity, we ignore all issues associated with the construction of innovations in practice and assume that true innovations are available to the researcher. Consider again the basic regression of interest:

$$Y_t = \beta X_t + \nu_t,$$

where X_t is correlated with ν_t ($\theta \neq 0$):

$$\nu_t = \theta X_t + \xi_t = \theta(E_{t-1}X_t + \epsilon_t) + \xi_t,$$

and where we use the definition of the innovation to X_t as

$$\epsilon_t = X_t - E_{t-1}X_t.$$

Rewriting the basic regression:

$$Y_t = \beta\epsilon_t + \underbrace{(\beta + \theta)E_{t-1}X_t + \theta\epsilon_t + \xi_t}_{\text{New Error } \chi_t}.$$

Obviously the regression of Y_t on the innovation to X_t, ϵ_t, is biased since

$$COV(\epsilon, \chi) = \theta VAR(\epsilon) \neq 0.$$

Conclusion: Innovations of an endogenous regressor are typically endogenous.

4.2 Innovations Lead to New Biases?

While replacing a regressor with innovations does not overcome the endogeneity problem, an important question is whether this is an otherwise innocuous procedure. We now show that it is not: using innovations in an unbiased regression typically generates an additional bias. Timing assumptions generally do not overcome this bias.

We therefore adopt the realistic timing assumptions (also apparently used in CRK), that current benefits depend on the past unemployment rate and that the current unemployment rate is only learned next period.

$$Y_t = \beta X_t + \nu_t,$$

where Y_t is unemployment and X_t are benefits in period t. Benefits in period t depend deterministically on unemployment in period $t-1$,

$$X_t = \gamma Y_{t-1}.$$

However when setting benefits in period $t-1$, benefits up to period $t-1$ are known but only the unemployment rate in period $t-2$ is known, so that the innovation in benefits is

$$\epsilon_t = X_t - E_{t-1}X_t = \gamma(Y_{t-1} - E_{t-1}Y_{t-1}) = \gamma(\nu_{t-1} - \nu_{t-2}).$$

Using ϵ as a regressor generates an additional bias if

$$COV(\gamma(\nu_{t-1} - \nu_{t-2}), \nu_t) \neq 0.$$

One such additional bias would arise, for example, when

$$COV(\nu_{t-1}, \nu_t) \neq 0,$$

i.e. unemployment is not i.i.d.[2]

Conclusion: Using innovations typically generates additional bias(es).

Two additional comments are worth making here.

1. The derivations above ignore the fact that even more biases are induced by the procedure used to construct innovations as discussed in HMM (2016b). Despite the claim to the contrary in CRK (2016b), CRK (2016a) make it clear that the construction of innovations conditions on revised unemployment. Thus, the lhs variable in the key regression is used to construct the rhs one. To appreciate the seriousness of these biases, just recall CRK (2016b)'s example discussed above in Section 3.1.[3]

[2] More generally suppose that the unemployment shock process is AR(1), $\nu_t = \rho\nu_{t-1} + \zeta_t$, then

$$\begin{aligned} COV((\nu_{t-1} - \nu_{t-2}), \nu_t) &= COV((\rho-1)\nu_{t-2} + \zeta_{t-1}, \rho^2\nu_{t-2} + \zeta_t + \rho\zeta_{t-1}) \\ &= \rho^2(\rho-1)VAR(\nu) + \rho VAR(\zeta) = VAR(\zeta)\frac{\rho}{1+\rho}, \end{aligned}$$

which is positive iff $\rho > 0$. Note that if the unemployment was actually i.i.d., then there would be no endogeneity problem and regressing unemployment on benefits would yield unbiased (large) estimates.

[3] HMM (2016a,b) also highlight the importance of CRK's decision to construct different transition matrices for different regions of revised unemployment. To get a sense of what this does, suppose there are separate transition matrices above and below the 6% unemployment rate. When unemployment is just below 6%, it is higher than the expectation of the transition matrix that is below 6%, so the innovation is positive. When unemployment rises to just above 6% the innovation all of a sudden becomes negative. Note that the benefit error does not need to change at all when the innovation changes (innovations change even when both revised and real time unemployment cross 6% at the same time). HMM (2016 a,b) illustrated this with an example of Nevada and California, where there was never a benefit error, yet benefit error innovations move over time. CRK (2016b) responded that those movements are not very large. But this is not the point. Such movements of innovations are pervasive in the CRK data, leading to further questions: (1) how much noise is introduced into the innovation series leading to the attenuation bias in their regression, and (2) whether those errors co-move

2. The robustness analysis in CRK (2016a) does not address any of these biases. First, including the measurement error in unemployment as a regressor instead of using it as an instrument is not helpful for correcting for the endogeneity bias or testing for it. Indeed, under the fundamental identification assumption, which we have not yet challenged, measurement error in unemployment is orthogonal to revised unemployment. Including random noise as a regressor cannot help overcome the endogeneity bias in other variables. Second, they include the lags of the lhs variable Y_t on the rhs and estimate coefficients on them. This also does not address the endogeneity or any other biases that we pointed out.

5 Some Empirical Questions about the Breakthrough

Of course, the theoretical discussion above was just a simple illustration that may or may not apply in practice. Unfortunately, CRK provide no alternative theoretical guidance leaving us with no choice but to investigate the validity of their identification arguments empirically. This allows us to see if the endogeneity problems outlined above are borne our empirically, and to see if there are other potentially viable identification strategies exploiting unemployment errors. The analysis in this section uses exclusively CRK's data. We adhere to their practice of reporting 90% confidence intervals throughout. We also follow CRK's standards when clustering standard errors. They cluster at the state and month level which tends to inflate standard errors. The significant results we report below hold despite this.

5.1 Innovations are Predictable

Consider estimating main regression in CRK using leads or lags of benefit error innovations

$$\tilde{u}_{s,t} = \beta_k \epsilon_{s,t+k} + \delta_s + \delta_t + \nu^k_{s,t}, \tag{7}$$

where $k \in \{-8, ..., 0, ..., 8\}$. The estimated $\hat{\beta}_k$ are plotted in Figure 1.

The left hand side of this figure, for $k \leq 0$, is reported in CRK and represents their main finding. They refer to it as the impulse response of unemployment to benefit error innovations. However, the right hand side of this figure is no less exciting. It shows that future benefit error innovations significantly predict current unemployment. The existence of such a pre-existing

systematically with revised unemployment leading to other biases of unknown sign and magnitude. In light of this, CRK (2016b)'s comment that "the criticisms raised by HMM simply do not apply to our work" seems misplaced. Questions like this are about the basic properties of CRK's estimator and identification strategy, and they must be answered for the results to be interpretable.

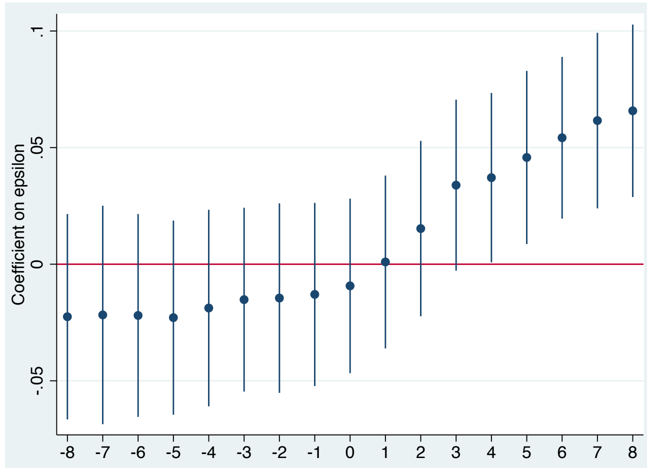


Figure 1: Future Benefit Error Innovations Predict Current Unemployment.

Note: Coefficients β_k and 90% ci from regression (7). According to CRK: Blue dots should lie on the red line $\forall k > 0$. As they do not, CRK's innovations are either endogenous with respect to past unemployment or economic agents can foresee the "innovations" and adjust their behavior prior to "innovations" occurring.

trend is generally considered problematics for attempts at causal inference using regression discontinuity or event study approaches, such as CRK's. But setting the econometric issues aside, the economics is really interesting.

Standard economic theory suggests that agents react in advance of anticipated policy changes. Thus, expecting a future increase in benefit error "innovation," job creation, which is a forward looking investment decision, falls today resulting in a lower job-finding rate and higher unemployment.[4] The adjustment is quite fast, at least in the standard search model, implying that much of the adjustment of job creation (and of unemployment in a standard search model) would have occurred prior to the actual change in policy. At the minimum, the basic optimality of firms' decisions implies no discrete jumps in vacancy posting at the time that expected policy changes are implemented. If "innovations" were anticipated by agents, then this logic would explain the finding that unemployment rises in response to future "innovations" but effectively stops rising by the time "innovations" actually occur. Thus, one interpretation of the pre-trend is that the econometrician-identified "innovations" were, in fact, predicted and acted upon in advance by economic agents. This interpretation reinforces the key lesson of Hagedorn et al. (2013) who argued that to obtain unbiased and interpretable estimates of the effects of unemployment benefits on the aggregate labor market, it is important to take into account that firms' job creation decisions depend on expectations of future policies.

[4]Job creation declines in response to an increase in benefit generosity because wages rise for a given level of workers' productivity. The logic is standard: An improvement of workers' outside option leads to a wage increase to prevent shirking in the efficiency wage model and directly affects the wage bargain in the search model.

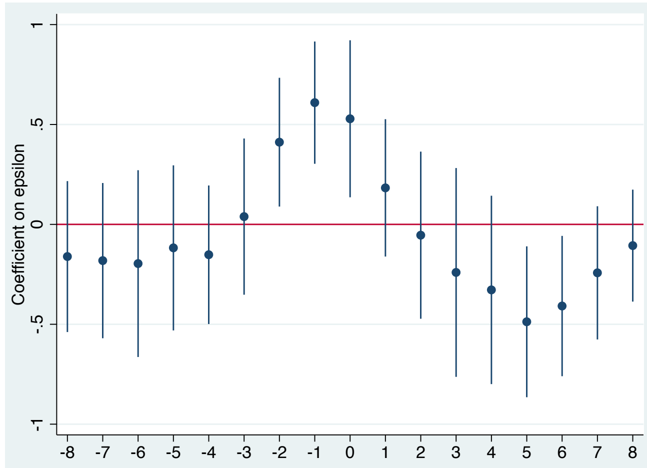


Figure 2: Past and Future Benefit Errors Innovations Predict Fraction Claiming UI.

Note: Coefficients β_k and 90% ci from regression (8). According to CRK: Blue dots should lie on the red line $\forall k > 0$. In the data: clear pre-trend in the fraction of unemployed claiming UI which responds significantly to future benefit error "innovations."

Of course, this evidence is also consistent with the presence of the mechanical endogeneity bias that is not overcome by the construction of benefit error "innovations." Specifically, we cannot tell apart whether (1) unemployment rises today in expectation of higher future benefit "innovations" and (2) benefit "innovations" occur in response to higher lagged unemployment.

These results should provide some comfort to those concerned that CRK's benefit error innovations are essentially noise because there is at least some informational content left in the series. CRK (2016a) make a similar case by plotting the impulse response of the fraction of the unemployed claiming UI, $\phi_{s,t}$, to a benefit error innovation. We replicate their analysis, but report a complete figure, including the coefficients on both leads and lags:

$$\phi_{s,t} = \beta_k \epsilon_{s,t+k} + \delta_s + \delta_t + \nu^k_{s,t}. \tag{8}$$

The results plotted in Figure 2 feature the same clear pre-trend in the fraction claiming UI as in unemployment in Figure 1. In particular, the fraction of unemployed claiming UI responds significantly to future benefit error "innovations." Quantitatively, the response of the fraction claiming UI prior to the "innovation" is nearly identical in absolute value to the response after the "innovation" occurs.[5]

[5]The economics behind these patterns is less clear. One possibility is that individuals who are currently not claiming benefits (and save them to be claimed in the future) start claiming when they expect an "innovation" a few months into the future that would lead to a cut in extended benefits.

5.2 Obvious Bias when using Benefit Error Innovations

In HMM (2016a,b) we pointed out that the bias in CRK's regression based on benefit error innovations, i.e., Eq. (6) is immediately evident in the data. Specifically, we documented that benefit error innovations, $\epsilon_{s,t}$, co-move negatively with hypothetical benefit duration, $\tilde{T}_{s,t}$, which, in turn, co-moves positively with the revised unemployment rate, $\tilde{u}_{s,t}$. Taken together, these findings imply that the error term (shocks that affect revised unemployment) in regression (6) is negatively correlated with benefit error innovations, leading to a downward bias in the estimated β.

We emphasize that these results were reported in HMM (2016a) on p. 10 and HMM (2016b) on p. 11 (slide "Evidence of the Bias in the Data"). They directly imply that CRK's benefit error innovations inherit the endogeneity problem of the benefit errors. CRK(2016b,c) have not acknowledged or contested this evidence. Moreover, the wording of their replies suggests that HMM (2016a,b) only discuss the endogeneity of benefit errors and do not show endogeneity of benefit error innovations. This is simply not true. Establishing the endogeneity of benefit errors was the crucial insight of HMM (2016a)—the endogeneity of the innovations to benefit errors generally follows as a corollary (see, e.g., Section 4.1 above). Nevertheless, we took the effort to specifically document the endogeneity of benefit error innovations constructed by CRK in the data.

In the next section we provide additional evidence that innovations are an endogenous regressor.

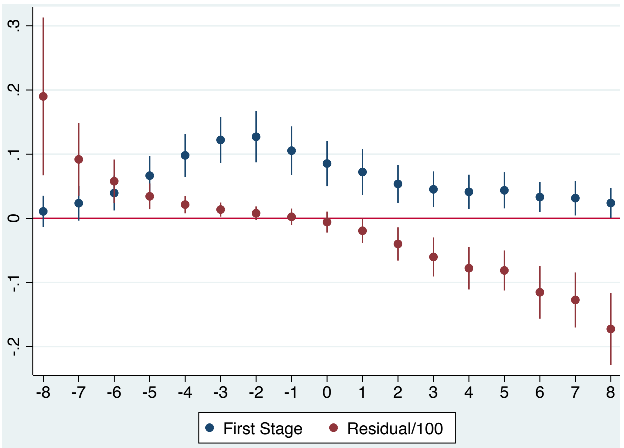
5.3 Innovations are Endogenous?

A standard Hausman test can be used to assess whether benefit error innovations are endogenous in CRK's benchmark regression (6). The test exploits the fundamental identification assumption in CRK (2016a) that $\hat{u}_{s,t}$ – the difference between real time and revised unemployment – is exogenous measurement error that is independent of revised unemployment, $\tilde{u}_{s,t}$. It is the measurement error in unemployment, however that must be inducing errors in benefits and innovations in benefit errors.

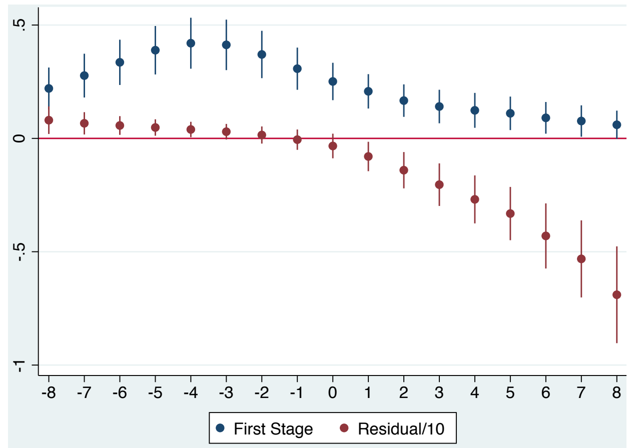
The blue dots in Figures 3(a) (for those looking at the document in black and white, these are the dots that always lie above the zero line in each Panel of Figure 3) correspond to estimated parameter γ_k from the following regression

$$\epsilon_{s,t} = \gamma_k \hat{u}_{s,t+k} + \delta_s + \delta_t + \nu^k_{s,t}, \tag{9}$$

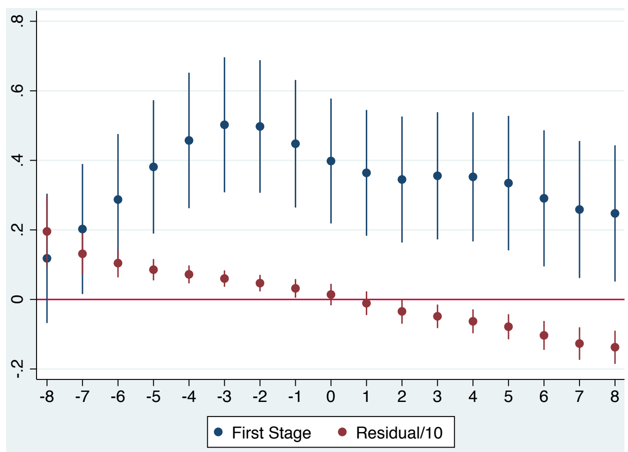
where $k \in \{-8, ..., 0, ..., 8\}$. Similarly, the blue dots in Figures 3(b) and 3(c) plot the estimated



(a) Endogeneity of Benefit Error Innovations



(b) Endogeneity of Benefit Errors



(c) Endogeneity of Real-Time Benefits

Figure 3: Hausman Test Reveals Endogeneity of Benefit Errors and Benefit Error Innovations.

Note: Coefficients β_k and 90% ci from regression (9) – blue dots, and regression (10) – red dots. According to CRK: Red dots should lie on the red line $\forall k$. In the data: each red dot deviating significantly from red line marks a failure of the exogeneity test for the benefit variable in the title of each panel. Moving from benefits to benefit errors to benefit error innovations does not resolve the endogeneity problem.

coefficients γ_k form the same regression but with benefit errors, $\hat{T}_{s,t}$, or benefit levels, $T_{s,t}$, as the lhs variable, respectively.

Note that these are the standard first stage regression that reveal the strength of $\hat{u}_{s,t}$ if used as an instrument. We observe that most leads and lags of $\hat{u}_{s,t}$ are indeed strong instruments for benefits, benefit errors and benefit error innovations.

In the second step, we include the residuals, $\nu^k_{s,t}$, from each of the first stage regressions corresponding to different leads or lags of $\hat{u}_{s,t+k}$ into the baseline regression in CRK:

$$\tilde{u}_{s,t} = \beta_\epsilon \epsilon_{s,t} + \beta^\nu_k \nu^k_{s,t} + \delta_s + \delta_t + \mu^k_{s,t}. \tag{10}$$

The estimated coefficients $\hat{\beta}_k^{\nu}$ are plotted in red in figure Figure 3(a). The same procedure is repeated for benefit errors rather than benefit error innovations in Figure 3(b) and for benefit levels in Figure 3(c).

Significant estimated coefficients $\hat{\beta}_k^{\nu}$ for most leads and lags imply the endogeneity of benefit error innovations and other benefit variables. In particular, this indicates that moving from benefits to benefit errors to benefit error innovations has not resolved the mechanical endogeneity problem in CRK's analysis.

We should emphasize that the informativeness of the Hausman test is conditional on unemployment errors being not only a strong, but also a valid instrument. In other words, the test relies on the fundamental identifying assumption in CRK (2016a) being correct. We will present evidence below implying that this is not the case. The failure of this identifying assumption, in addition to the limited power of the test, may help explain the puzzling rejection of the endogeneity of real-time benefits when contemporaneous unemployment errors are used as an instrument, see Figure 3(c).

5.4 Root of the Problem: Identification Assumptions are Flawed

The argument of CRK (2016a) can be summarized in four steps.

1. The fundamental identification assumption is that $\hat{u}_{s,t}$, the measurement error in unemployment, is uncorrelated with the true revised unemployment $\tilde{u}_{s,t}$.

2. Measurement errors in unemployment, $\hat{u}_{s,t}$, induce errors in benefits $\hat{T}_{s,t}$ which are uncorrelated with hypothetical benefits $\tilde{T}_{s,t}$.

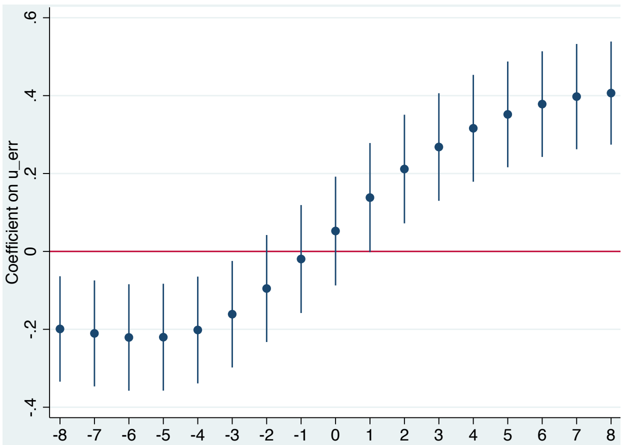
3. Benefit error innovations, $\epsilon_{s,t}$, are exogenous and unpredictable.

4. Benefit error innovations have no impact on macroeconomic variables such as unemployment.

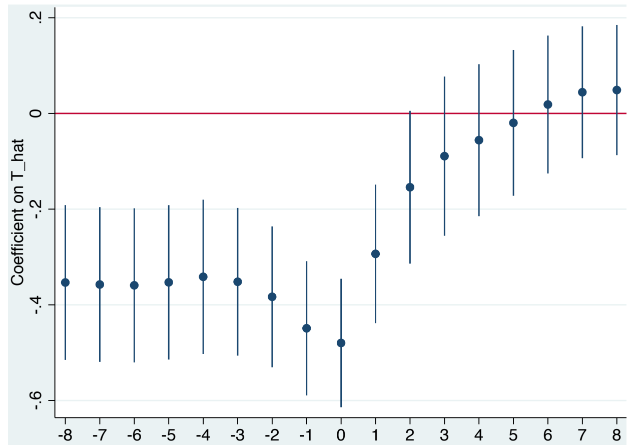
If correct, this argument implies tight restrictions on the relationship between several key variables in the data. In particular, all dots in each panel of Figure 4 should be at zero while they are not in the data. These failures are informative regarding the flaws in the logic of CRK's argument.

Consider first Figure 4(a) which plots estimated coefficients β_k from the regressions

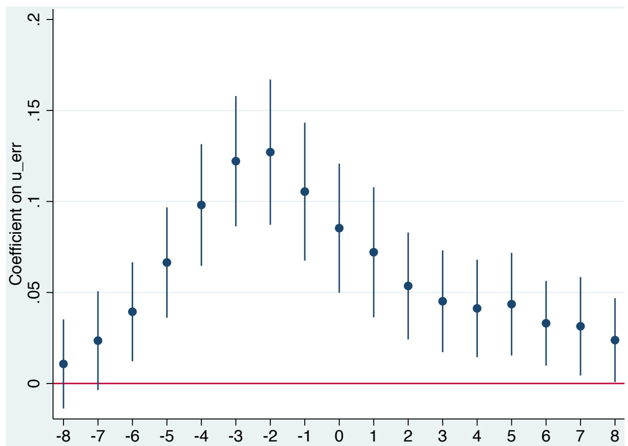
$$\tilde{u}_{s,t} = \beta_k \hat{u}_{s,t+k} + \delta_s + \delta_t + \nu_{s,t}^k, \tag{11}$$



(a) Revised Unemployment depends on Unemployment Errors



(b) Hypothetical Benefits depend on Benefit Errors



(c) Benefit Error Innovations depend on Unemployment Errors

Figure 4: The Failure of CRK's Identifying Assumptions.

Note: Coefficients β_k and 90% ci from reg. (11) – Panel 4(a), reg. (12) – Panel 4(b), and reg. (13) – Panel 4(c). According to CRK: Blue dots should lie on the red line $\forall k$. In the data: Most blue dots are statistically different from the red line, indicating dramatic failure of CRK's identifying assumptions.

for $k \in \{-8, ..., 0, ..., 8\}$. The fact that leads and lags of "measurement error" in unemployment predict revised unemployment, points to the failure of the fundamental identification assumption on which CRK (2016a) is based. This failure is, perhaps, not surprising, because unemployment data revisions not only reflect the use of better data, but systematic methodological changes.[6]

[6]The fact that lagged unemployment errors predict current revised unemployment could also be induced by the failure of the last step of CRK's argument. Specifically, an increase in unemployment error, say due to the rise of the real-time unemployment, leads to an increase in real-time benefits, leading to a significant increase in revised unemployment. To assess this possibility, we can control for (lags of) real-time benefits in the regression:

$$\tilde{u}_{s,t} = \underset{(0.070)}{-0.301}\, \hat{u}_{s,t-3} + \underset{(0.027)}{0.198}\, T_{s,t} - \underset{(0.030)}{0.001}\, T_{s,t-1} + \underset{(0.030)}{0.007}\, T_{s,t-2} + \underset{(0.023)}{0.098}\, T_{s,t-3} + \delta_s + \delta_t + \nu_{s,t}.$$

Thus, we have established that errors in unemployment are not exogenous "measurement errors." This implies that errors in benefits are not "measurement errors" either. To see this directly in the data, consider Figure 4(b), which plots estimated coefficients β_k from the regressions

$$\tilde{T}_{s,t} = \beta_k \hat{T}_{s,t+k} + \delta_s + \delta_t + \nu_{s,t}^k. \tag{12}$$

The fact that leads and lags of "measurement error" in benefits predict hypothetical benefits, is expected given the failure of the fundamental identification assumption.

Finally, if benefit errors are not "measurement error," benefit error innovations are not measurement error either. This can be also seen directly in the data. In Figure 4(c) we plot estimated coefficients β_k from the following regressions

$$\epsilon_{s,t} = \beta_k \hat{u}_{s,t+k} + \delta_s + \delta_t + \nu_{s,t}^k. \tag{13}$$

As can be by now expected, benefit error "innovations" are predicted by leads and lags of unemployment errors.

5.5 Failure of Fundamental Identification Assumptions also Invalidates the IV Strategy

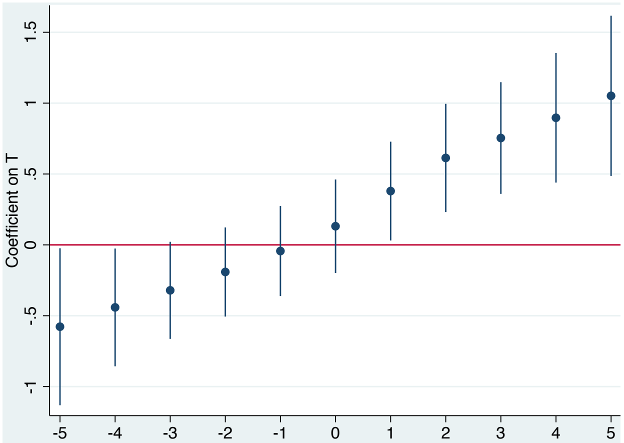
In HMM (2016b) we proposed that if the fundamental identifying assumption of CRK is satisfied, a preferred approach to inference is through the IV regression that instruments benefits $T_{s,t}$ in Eq. (1) with unemployment errors, $\hat{u}_{s,t}$. Although not necessary, one can also use $\hat{u}_{s,t}$ as an instrument for benefit errors, $\hat{T}_{s,t}$, in Eq. (4) or benefit error innovations, $\epsilon_{s,t}$, in Eq. (6). As we have established in Section 5.3 by estimating the first stage regressions, the leads and lags of $\hat{u}_{s,t}$ are very strong instruments for all benefit variables (see Figure 3). In addition to the strong first stage, leads and lags of unemployment errors $\hat{u}_{s,t}$ are independent of revised unemployment, $\tilde{u}_{s,t}$, by the fundamental identifying assumption in CRK (2016a). Thus, they can be used as instruments to overcome the endogeneity bias in the corresponding regressions.

In Figure 5 we report the coefficients on benefit variables from regressions (1), (4), and (6), where $\hat{u}_{s,t+k}$, for $k \in \{-5, ..., 0, ..., 5\}$ is used as an instrument for $T_{s,t}$, $\hat{T}_{s,t}$, and $\epsilon_{s,t}$, respectively.[7]

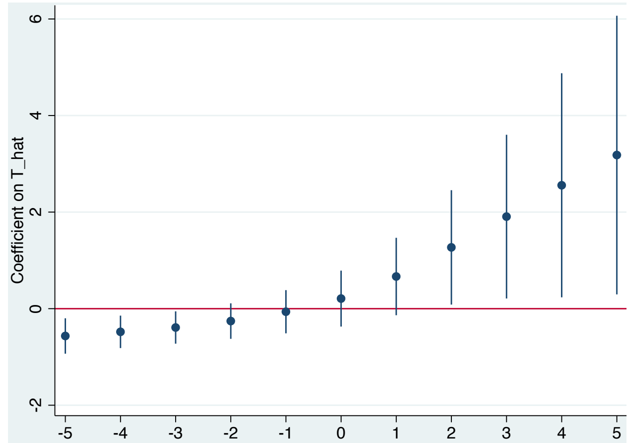
The results are striking. In particular, the magnitude of the coefficients make no economic sense. For example, in HMM (2016b) we show that a coefficient of 0.573 for $k = 0$ in Figure

The results are supportive of the presence of this mechanism, although it is not sufficient to explain the failure of the fundamental identification assumption.

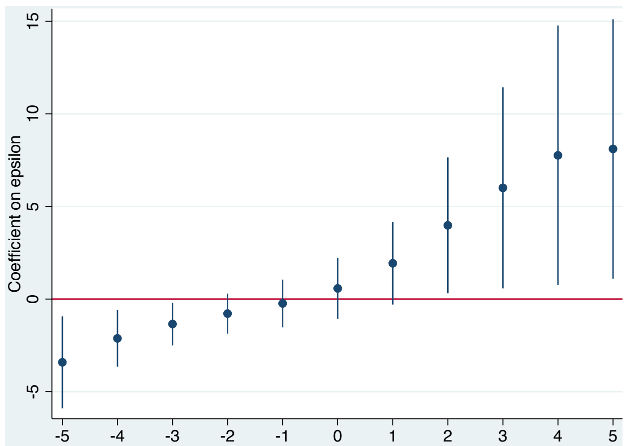
[7]The estimates and their standard errors for other leads and lags can be seen in the log file accompanying this Note. These coefficients become so large that the figure becomes unreadable if they are included.



(a) Revised Unemployment on Real Time Benefits



(b) Revised Unemployment on Benefit Errors



(c) Revised Unempl. on Benefit Error Innovations

Figure 5: Nonsensical results using unemployment errors as instruments.

Note: Coefficients β and 90% ci from from regressions (1), (4), and (6), where $\hat{u}_{s,t+k}$, for $k \in \{-5, ..., 0, ..., 5\}$ is used as an instrument for $T_{s,t}$ – Panel 5(a), $\hat{T}_{s,t}$ – Panel 5(b), and $\epsilon_{s,t}$ – Panel 5(c), respectively. Switching signs across leads and lags in the same panel is similar to the results based on CRK's benefit error innovations – see Fig. 1 above. Failure of both estimators is due to the failure of CRK's fundamental identifying assumption documented in Fig. 4.

5(c) implies that a one month benefit error innovation leads to an increase of unemployment by 175 percentage points. Moreover, even if this particular coefficient is statistically insignificant, many even larger ones are.

In HMM (2016b) we reported just the estimates for $k = 0$ and did not cluster the standard errors.[8] CRK (2016c) refer to such an estimate in Figure 5(a) and criticize the IV approach by noting:

> ... we repeat the IV regression but forwarding the dependent variable by one period. The estimated coefficient reverses sign, which makes little economic sense, but

[8] The only point of the slide was that the IV regressions yield implausibly large coefficients, and clustering s.e. was not relevant for that argument.

remains statistically insignificant when properly adjusting for the standard errors.

CRK (2016c)'s critique only mentions the coefficient from using one lag but not what happens if they use one or more leads (which are strong and valid instruments). As can be seen in Figure 5(a), this implies positive and statistically significant effects of benefit extensions on unemployment.[9]

We agree with CRK (2016c) that the fact that the estimated coefficient in IV regressions reverses sign when moving from leads to lags "makes little economic sense." However, the same sign reversal is evident in Figure 1 reporting the coefficients from CRK's main regression based on benefit error innovations. Thus, both estimators are similarly flawed. The reasons for this have nothing to do with the linearity of the the first stage in the IV regressions (we verified that the results are robust to using a nonlinear first stage), or with the serial correlation of benefit errors and clustering procedures (these are simply irrelevant). The true reason is that the fundamental identification assumption underlying CRK's research strategy fails—as we have shown in the preceding Section.

Given the overwhelming importance of the research question and the intriguing results in CRK (2016a), we spent a long time thinking about their research strategy. Indeed, we are very excited when other researchers provide novel identification arguments allowing them to exploit new sources of variation to measure the effect of unemployment benefit extensions on labor market outcomes. Given the fact that CRK's strategy of using innovations does not overcome the endogeneity problem, we were hoping that we could help them develop a robust alternative measurement strategy built solely on their fundamental identifying assumption. This would have been clean, transparent, and informative. However, we discovered that the fundamental identifying assumption itself is not satisfied in the data, suggesting this is not a promising strategy going forward. Fortunately, our discussion of the literature in HMM (2016a) suggests a number of alternative strategies that appear more promising.

[9] We here use that, according to CRK's main assumption, the unemployment error is a strong instrument at various leads and lags whereas CRK (2016c) shift both the lhs variable and the instrument at the same time. Conducting CRK (2016c)'s less informative experiment at the same leads and lags as in Figure 5(a) does not change our conclusion.

References

Chodorow-Reich, G. and L. Karabarbounis (2016a): "The Limited Macroeconomic Effects of Unemployment Benefit Extensions," Working Paper 22163, National Bureau of Economic Research.

——— (2016b): "Reply to Hagedorn, Manovskii, and Mitman on "Interpreting Recent Quasi-Experimental Evidence on the Effects of Unemployment Benefit Extensions," Discussion Paper, July 17, Harvard University and University of Minnesota.

——— (2016c): "Reply to Manovskii's Discussion on "The Limited Macroeconomic Effects of Unemployment Benefit Extensions," Discussion Paper, July 28, Harvard University and University of Minnesota.

Hagedorn, M., F. Karahan, I. Manovskii, and K. Mitman (2013): "Unemployment Benefits and Unemployment in the Great Recession: The Role of Macro Effects," National Bureau of Economic Research Working Paper 19499.

Hagedorn, M., I. Manovskii, and K. Mitman (2016a): "Interpreting Recent Quasi-Experimental Evidence on the Effects of Unemployment Benefit Extensions," National Bureau of Economic Research Working Paper 22280.

——— (2016b): "Discussion of "The Limited Macroeconomic Effects of Unemployment Benefit Extensions" by Gabriel Chodorow-Reich and Loukas Karabarbounis," NBER EF&G Meeting, July 16.