RAND

The Relationship Between the Economy and the Welfare Caseload: A Dynamic Approach

Steven Haider, Jacob Klerman, and Elizabeth Roth

DRU-3003

February 2002

Labor and Population Program

Working Paper Series 03–06

THE RELATIONSHIP BETWEEN THE ECONOMY AND THE WELFARE CASELOAD: A DYNAMIC APPROACH

STEVEN HAIDER, JACOB KLERMAN, AND ELIZABETH ROTH
RAND
1700 Main Street
Santa Monica, CA 90407

First Draft: June 2000
Last Revision: February 2002

Address correspondences to Steven Haider (sjhaider@rand.org). This work is funded by U.S. Department of Health and Human Services (#99ASPE34A). The opinions and conclusions expressed herein are solely those of the authors and should not be construed as representing the opinions or policy of any of the sponsors or related agencies. Some of the analyses use California administrative data that cannot be released to other researchers without the permission of the State of California. We gratefully acknowledge the expert programming assistance of Janet Hanley and Laurie McDonald.

ABSTRACT

Nationally, the welfare caseload declined by more than fifty percent between 1994 and 2000. Considerable research has been devoted to understanding what caused this decline. Much of the literature examining these changes has modeled the total caseload (the stock) directly. Klerman and Haider (2001) model the underlying flows and show analytically and empirically that previous methods are likely to be biased because they ignore important dynamics. However, due to their focus on the bias of the stock models, they present only limited results concerning the robustness of their findings and utilize only a single measure of economic conditions, the unemployment rate. This paper examines the robustness of the basic stock-flow model developed in Klerman and Haider (2001), considering both richer dynamic specifications and richer measures of economic condition. We find that more complex dynamic specifications do not change the substantive conclusions, but richer measures of the economy do. While a model that only includes the unemployment rate attributes about half of the California caseload decline between 1995 and 1998 to the economy, models that incorporate richer measures of the economy attribute more than ninety percent of the decline to the economy.

Nationally, the welfare caseload declined by more than fifty percent between 1994 and 2000. The causes of this large decline are the subject of considerable debate. The decline occurred simultaneously with a robust economic expansion, a series of major welfare reforms, and major changes in several other public policies such as the federal Earned Income Tax Credit (EITC) and the minimum wage. Understanding the separate roles of each of these possible causes is important for evaluating the effectiveness of welfare reform and planning for future budgetary outlays.

Many studies have considered why the caseload declined. Most of the studies model the change in the caseload directly, reaching widely varying conclusions (e.g., CEA, 1997; Levine and Whitmore, 1998; CEA, 1999; and Ziliak et al., 2000). Klerman and Haider (2001) argue that much of the variation in conclusions across these studies can be explained with an explicit model of the underlying caseload dynamics. Specifically, they develop a model that views today's caseload as the result of previous flows on and off of welfare. They demonstrate that the conventional methods that model the caseload stock are likely to be biased. Furthermore, they develop an alternative strategy based on estimating the flow relationships and then simulated the changes on the caseload level (the "stock"). They implement this stock-flow approach on California data and find that approximately half of California's caseload decline from its peak in 1995 to the end of their data in 1998 can be attributed to the improving economy. Because that paper focused on the biases associated with modeling the stocks directly, they only present limited results on the robustness of their findings.

In this paper, we examine in greater detail the specification of a stock-flow model and the measurement of economic conditions. First, we examine the sensitivity of the results to richer specifications of the underlying dynamics, including allowing the process of re-entry to differ from initial entry and allowing the effect of the economy to vary with duration. Second, much of the previous literature assessing the change in the caseload uses only the unemployment rate to measure economic conditions. Two recent studies suggest that a richer set of measures of the economy substantially improves the fit of the model and perhaps increases the share of observed changes due to the economy (see

Bartik and Eberts, 1999, and Hoynes, 2000). We also estimate dynamic models with additional measures of the economy.

Our results suggest two main conclusions. First, we demonstrate that the simple dynamic structure estimated in Klerman and Haider (2001) is sufficient to capture empirically the underlying dynamic process of how the economy affects the welfare caseload. Second, including measures of earnings in the retail trade sector substantially increases the estimated role for the economy in explaining the caseload decline: our preferred model suggests that approximately ninety percent of the caseload decline in California is due to the improving economy, broadly defined. We interpret this finding to imply that the conventional unemployment rate is an imperfect proxy for economic conditions for the welfare population. Taken together, these findings provide further support for the conclusion that conventional stock models underestimate the role of the improving economy for the caseload decline. It is worth noting, however, that welfare reform occurred later and was less drastic in California than in most other states. Thus, because the role of policy might have been larger in other states, it is possible that the role of the economy would be smaller if a similar approach could be applied.

The balance of this paper is organized as follows. We begin with a brief review of the literature. The second section discusses our data. The third section describes the baseline model and findings. The fourth section explores the sensitivity of these findings to altering the dynamic structure of the model. The fifth section considers the sensitivity of the model to using additional measures of economic conditions. We conclude in section six with a summary and discussion of the findings.

1. A REVIEW OF THE LITERATURE

Most previous studies have directly modeled the aggregate welfare caseload to assess its various determinants. In particular, previous studies have estimated models of the form,

$$(1) \qquad \ln n_{jt} = \alpha + \beta Y_{jt} + \lambda W_{jt} + \pi X_{jt} + \gamma_t + \delta_j + \phi_j t + \varepsilon_{jt}$$

where, for state j and year t, n_{jt} is the per capita welfare caseload, Y_{jt} is the unemployment rate, W_{jt} is measures of the welfare policy, and X_{jt} is a vector of other regressors. Some papers have augmented this model to include lagged values of the unemployment rate, lagged values of the dependent variable, and other control variables (see CEA, 1997; Levine and Whitmore, 1998; CEA, 1999; Figlio and Ziliak, 1999; and Ziliak et al, 2000).[1]

These studies have come to widely varying conclusions. CEA (1997) attributes 44 percent of the decline in the caseload from 1993 to 1996 to economic conditions and 31 percent of the decline to welfare waivers. Ziliak, et al (2000), examining the same question, attributes nearly two-thirds of the decline to economic conditions and nothing to welfare waivers. Figlio and Ziliak (1999) attempt to reconcile these results and conclude that the primary reason for different results rests with differences in modeling the "dynamics," where dynamics refers to whether lagged dependent variables are included and whether the models are estimated in levels or differences. CEA (1999) updates the CEA (1997) and also notes that their results are highly sensitive; in particular, CEA (1999) reports that the estimated role of the economy is reduced by one half when the regression model includes two annual lags of the unemployment rate versus a regression model with only one annual lag of the unemployment rate.

This literature has developed nearly independently of a body of research that directly examines flows onto and off of welfare (e.g., Hutchens, 1981; Bane and Ellwood, 1986; and Hoynes, 2000).[2] Those studies provide strong empirical evidence that welfare receipt in a particular period is dependent on previous welfare receipt and length of welfare receipt, even conditional on covariates.

[1] Other studies have used similar models to examine the impact of welfare reform on other outcomes (Moffitt, 1999; Schoeni and Blank, 2001; Currie and Grogger, 2000).

[2] Hutchens (1981) examines separately entry onto and exit off of welfare. Bane and Ellwood (1986) and Blank (1989) examine the determinants of welfare spell lengths. Moffitt (1992) provides a useful review of the earlier studies. Hoynes and MaCurdy (1994) examine how changes in program generosity affect spell length. Gittleman (2000) and Hoynes (2000) are more recent studies of welfare dynamics.

Klerman and Haider (2001) reconsider the caseload literature by specifying a model that is directly based on the underlying flows onto and off of welfare and then simulating the impact on the caseload levels. They refer to this model as a stock-flow model. Relying on such a model, they demonstrate analytically that the conventional models are likely to be mis-specified whenever welfare receipt in a particular period depends on welfare receipt in previous periods. They then demonstrate that this misspecification can explain several empirical anomalies in the earlier national caseload literature.

Klerman and Haider (2001) also report new estimates of the role of the economy that are based directly on this stock-flow model. The model is a direct application of Markov chain models that has been used in many other contexts (e.g., Heckman and Walker, 1992; and Moffitt and Rendall, 1995). To implement the model, they first estimate the relationship between the flows onto and off of welfare and the unemployment rate. Given the parameter estimates and the initial distribution of the population on and off aid, they then simulate the implied path of the caseload for an arbitrary path of the explanatory variables. Klerman and Haider (2001) conclude that about half of the California caseload decline can be attributed to changing economic conditions, as measured by the unemployment rate.

2. THE DATA

Our primary data set is based on micro-level, administrative data from the California Medi-Cal Eligibility Determination System (MEDS). We also rely on various publicly available data sets for population characteristics and measures of economic conditions.

The Medi-Cal Eligibility Data System (MEDS). Directly estimating the stock-flow model requires panel data on individuals, with sufficiently varying explanatory variables (e.g., economic conditions), for a sufficiently long time period, and for a sufficiently large sample. Although national data would be preferable, suitable data unfortunately do not appear to exist. There is no individual-level administrative data at the national level and the available panel surveys have samples too small to allow estimation of the

transitions that are the focus of this study (e.g., the Panel Study of Income Dynamics or the Survey of Income and Program Participation).[3]

California is a large state with more than twenty percent of the U.S. welfare caseload and more than ten percent of the U.S. population. Thus, caseload trends in California comprise a significant share of the caseload trends in the U.S. Moreover, there is significant diversity in economic conditions across California's counties. This diversity allows us to use an identification strategy that is similar to that used in the national literature. Whereas the national literature relies on cross-state variation (controlling for time fixed effects) to identify the role of the economy, we rely on cross-county variation (controlling for time fixed effects) to identify the role of the economy.

Using California administrative data has one major disadvantage: we do not have a plausible identification strategy to separate policy effects from more general statewide time effects.[4] Thus, while we can directly estimate the effect of the economy, we cannot directly estimate the effect of welfare reform or other statewide (or national) policy changes. We show below, however, that our methods substantially increase the estimated effect of the economy, and thus decrease the possible effect of welfare reform.

The MEDS provides a monthly roster of all Medi-Cal (Medicaid in California) participants in California from January 1987 to December 1998. The source of eligibility for Medi-Cal is also indicated, and because all welfare participants are eligible for Medi-Cal, the MEDS effectively serves as a roster of all welfare recipients in California.[5] Individual identifiers make it possible to follow individuals over time. In addition, the

[3] Hoynes (2000) argues that the small sample sizes usually available in survey data have caused researchers to conclude that economic conditions are unimportant. The paper relies on a different extract from the same database for its analysis.

[4] The national stock literature uses time fixed effects in their specifications, and we adopt a similar specification for our flow models. This allows us to isolate the impact of moving to a stock-flow model as compared to the conventional models in the literature. However, the major policy changes that happened in California during our time period mainly took place at the state level.

[5] Throughout this paper, we use the term "welfare" to refer to the Aid For Families with Dependent Children (AFDC) program that was changed to the Temporary Assistance for Needy Families (TANF) program. The programs provide financial assistance to needy families (usually headed by a single mother) with children. Program participation can be identified with the MEDS data because welfare recipients are categorically eligible for Medi-Cal (the California implementation of Medicaid), and such eligibility is designated in the database.

MEDS contains basic information about age, gender, type of aid, and race for each person on welfare. Because the MEDS data are recorded as part of an ongoing administrative process, biases associated with self-reports are absent.

We construct an analysis file by drawing a stratified random sample of approximately 3 percent of the individuals on the full MEDS file. Consistent with our focus on the effect of county level variation in the economy, the stratified sample is chosen to yield approximately equal number of observations in each of California's counties.[6] This scheme results in an analysis file that contains a sample of 282,381 people who received cash assistance, comprising 487,641 spells and 10,966,420 person-months, during the years 1989 to 1998 (our eventual sample period). We consider a person as the unit of analysis rather than a case because persons are well-defined longitudinally, even if case membership changes.

We modify this basic extract along two dimensions to obtain our analysis file. First, to avoid some problems related to small samples, we aggregate California's five smallest counties into a single "county" for analysis purposes. The five smallest counties are Alpine, Colusa, Modoc, Mono, and Sierra; combined, their welfare population represents well under one percent of the state's welfare population. We perform all of our analyses on these 53 counties and 1 county group; hereafter, we will simply refer to our analysis considering 54 counties. Second, previous research indicates that there is considerable "churning" on and off welfare in the MEDS data. This churning is likely due to underlying administrative procedures rather than "real" entrances and exits (see Hoynes, 2000). To mitigate such concerns, we recode one-month spells on and off of aid as not having occurred, following Hoynes (2000).

Other Data Sources. We also use data from several other sources. First, we present tabulations from California's official caseload filings with the California Department of

[6] Specifically, we draw a monthly random sample of individuals who are entering aid for the first time in our sample period. We then follow each of these individuals to the end of our sample period. This strategy effectively provides a sample that contains an initial cross-section of everyone on aid, and then refreshes this initial cross-section with a sample of new entrants.

Social Services, the so-called CA237 data. These data are used to verify the quality of the caseload information available in the MEDS. The CA237 data are described in detail in Haider, et al. (1999).

We use population estimates for each county in California from the Demographic Research Unit of the California Department of Finance.[7] These data are designated as the single official source of demographic data for State planning and budgeting. The data are only available annually, so we use simple linear interpolation to obtain monthly county data. Throughout, we consider the population at risk of going on aid to be all individuals aged 0 to 49.[8]

Our models are designed to explore the effects of the aggregate state of the economy on the aggregate welfare caseload. As one measure of the aggregate state of the economy, we use county-level unemployment estimates produced by the U.S. Department of Labor as part of their Local Area Unemployment Statistics program. These estimates are produced through a cooperative effort by federal and state agencies and are based on the Current Population Survey data, Current Employment Statistics data, and state unemployment insurance data.

To derive employment rates, we use the civilian employment estimates produced by the California Employment Development Department. We calculate the employment rates by dividing by the county population aged 15-64. The underlying data for these estimates come from the unemployment insurance filings (the so-called ES202 data).

Finally, we use county-level retail trade earnings data prepared by the U.S. Department of Commerce, Bureau of Economic Analysis and archived at University of Virginia Geospatial and Statistical Data Center.[9] We convert all earnings measures to January 1998 dollars using BLS's consumer price index for urban consumers. The data

[7] We obtained these data from their website (http://www.dof.ca.gov/html/demograp/data.htm) in December 2000.

[8] We include men and women because our analysis includes children and the smaller AFDC-Unemployed Parent program (AFDC-UP), which provides welfare benefits to two parent families in which the husband has recently lost a job.

[9] We obtained these data from their website (http://fisher.lib.virginia.edu/reis) in January 2001.

are also only available by year, so we use simple linear interpolation to obtain monthly county data.[10]

Descriptive Results. Figure 1 plots the aggregate monthly caseload estimated from our MEDS-based analysis file, the official state caseload counts (based on county-level CA237 reports), and the unemployment rate. The figure shows that the MEDS tracks the official CA237 caseload very well.

Figure 1 also demonstrates that the paths of the caseload and the unemployment rate suggest a role for the economy in explaining the caseload decline. In particular, the caseload increased during the early 1990s and then declined during the late 1990s, similar to the trend for the United States as a whole. At the peak of the welfare caseload in March 1995, there were approximately 2.7 million people receiving AFDC/TANF in California. In the last month of our sample period, December 1998, there are only 1.9 million people on aid, representing a decline from the peak of 31 percent. Because the population increased during the same time frame, the per capita caseload declined slightly more, about 33 percent. Turning to the unemployment rate, the figure shows that the unemployment rate increased then decreased, following a similar pattern to that of the caseload. In particular, the unemployment rate declined from 10 percent at its peak to 6 percent at the end of our sample period.

The basis of the stock-flow model is that entry onto and exit from welfare varies with economic conditions. Table 1 presents the average monthly entry rate and the average monthly continuation rate (i.e., one minus the exit rate) for two-year intervals between 1989 and 1997. These tabulations reveal several important characteristics of the data. First, the levels of the entry and continuation rates in Table 1 are quite different, suggesting that dynamics could be important. The average monthly entry rate for those

[10] As noted, some of our underlying data are at the monthly level (welfare recipients, unemployment, and employment) and some of our data are at the annual level (population and earnings). We choose to perform our empirical analysis at the monthly level, and thus interpolate the annual data to the monthly level. Another empirical strategy would be to aggregate the monthly data to the annual level for analysis purposes. We believe our strategy to be more appropriate because welfare receipt is a discrete time process that occurs at the monthly level. Moreover, given that we will include flexible time splines in all of our analysis, the mechanical time correlation implied by the interpolation should not affect our estimates.

who were not on welfare in 1989-90 was 0.0032 and average monthly continuation rate for those who were on welfare for 2 to 5 months was 0.938. Second, both the entry and continuation rates are counter-cyclical. The entry rate increased during the recession of the early 1990s (from 0.0032 to 0.0037) and then declined during the recovery (back to 0.0024). Similarly, for all durations, the continuation rate increased then decreased. Both the entry and continuation rate patterns would cause the aggregate caseload size to vary counter-cyclically with economic conditions. Finally, the continuation rate is higher for individuals who have been on aid longer (i.e., there is duration dependence).

3. AN EMPIRICALLY TRACTABLE STOCK-FLOW MODEL

In this section, we first describe the basic stock-flow model that was examined in Klerman and Haider (2001), which will be denoted as our baseline model. Second, we develop the methods for estimating the underlying flow relationships. Third, we present the results for the baseline model.

The Baseline Stock-Flow Model. The stock-flow representation is a straightforward application of a Markov Chain process. Specifically, consider a standard Markov Chain,

$$(2)\quad \underset{(Q\times 1)}{S_t} = \underset{(Q\times Q)}{M(X_t,\theta)}\,\underset{(Q\times 1)}{S_{t-1}},$$

where S_t is a vector that contains the number of individuals in each of Q states, M is the transition matrix between the states (that depends on regressors X_t and a parameter vector θ), and t indexes time. A simple specification would allow for only two states, where an individual is either on welfare or off welfare. A richer specification, consistent with the finding of duration dependence, would disaggregate the on-welfare state by length of the current spell. The effect of potential explanatory variables X_t is captured in the transition matrix M.

More specifically, let $e(Y_t)$ denote the probability that an individual enters aid (unconditional on duration) and $c^k(Y_t)$ denote the probability that a person who has been

on aid for k periods remains on aid, with both flow rates being a function of economic conditions, Y_t. Then, aid receipt can be represented with the equation,

$$(3)\quad \begin{bmatrix} S_{r,1,t} \\ S_{r,2,t} \\ S_{r,3,t} \\ \vdots \\ S_{r,\bar{k}-1,t} \\ S_{r,\bar{k},t} \\ S_{n,t} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & e(Y_t) \\ c^1(Y_t) & 0 & \cdots & 0 & 0 & 0 \\ 0 & c^2(Y_t) & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \cdots & c^{\bar{k}-1}(Y_t) & c^{\bar{k}}(Y_t) & 0 \\ 1-c^1(Y_t) & 1-c^2(Y_t) & \cdots & 1-c^{\bar{k}-1}(Y_t) & 1-c^{\bar{k}}(Y_t) & 1-e(Y_t) \end{bmatrix} \begin{bmatrix} S_{r,1,t-1} \\ S_{r,2,t-1} \\ S_{r,3,t-1} \\ \vdots \\ S_{r,\bar{k}-1,t-1} \\ S_{r,\bar{k},t-1} \\ S_{n,t-1} \end{bmatrix},$$

where S_{rkt} is the number of individuals who are receiving aid for the kth consecutive period at time t and S_{nt} is the number of individuals not on aid. To handle an initial conditions problem, this formulation assumes that the continuation rate varies through period $\bar{k}$ and is constant thereafter; we discuss this assumption in further detail below.

This equation can be used to simulate the impact of economic conditions on the caseload stock implied by the underlying flow relationships. We first estimate models for the flows (i.e., the entry rate and the continuation rate) to obtain estimates of the parameter vector θ. Then, for any arbitrary specification for economic conditions $\tilde{Y}_t$, we calculate the implied transition matrix M and simulate the caseload for the following period. Thus, given an initial stock S_t and any arbitrary path for economic conditions, $\left\{\tilde{Y}_{t+h}\right\}_{h=1}^{H}$, we can simulate the future stock in period $t+h$ as,

$$(4)\qquad S_{t+h} = \left(\prod_{p=1}^{h} M\left[\tilde{Y}_{t+p},\theta\right]\right) S_t .$$

To explore the effect of the economy on the welfare caseload, we simulate the model for the observed path of the economy and for an explicitly specified counterfactual path. The difference between the two paths of the total caseload is the implied effect of the economy.

We specify caseload simulations that are intended to provide answers to the question that is usually posed in the literature: "How much of the caseload decline (from its peak) is due to economic conditions?" To answer this question, we compare the results from a simulation based on the actual path of the economy and other factors (proxied for by the time fixed effects) to the results from a simulation where we instead specify that the unemployment rate path follows its actual path until its peak in January 1993 and then remains constant. The difference between the two paths is our estimate f the caseload decline that can be attributed to the improvement in the economy.

Specifying the Flow Relationships. The basis for the simulation model is the empirical relationship between the flows (the entry and continuation rates) and the economic conditions. Following much of the literature, we estimate the entry rate with a logit specification (e.g., Bane and Ellwood, 1986; Blank and Ruggles, 1996; and Hoynes, 2000). For the continuation rate, we employ a linear probability model for computational ease, given that our micro data set includes approximately 10 million person-months. Both empirical models are traditional hazard models that examine the probability that someone changes states (either on to or off of welfare) conditional on being at risk of changing states.

Because the MEDS data include information only for those on welfare, we estimate our model for the entry rate's dependence on the unemployment rate using a grouped-data logit model (Maddala, 1983). We calculate the entry rate for county j in month t, e_{jt}, as the ratio of the number of entrants observed in the MEDS relative to the number of people at risk of going on aid (i.e., the total population less those already on welfare). We then estimate a grouped-data logit model at the county-month level that includes the unemployment rate and fixed effects for time and county,

$$(5)\quad \ln \frac{e_{jt}}{1-e_{jt}} = \alpha_e + \beta_e Y_{jt} + \gamma_t + \delta_j + \varepsilon_{jt}.$$

Rather than including a full set of dummy variables for each calendar month, we include a discontinuous piecewise linear annual spline to capture a general time trends. In other

words, we include year dummy variables and the appropriate interactions to allow for different linear time trends by year, and we do not restrict the linear trends to be continuous between years (i.e., between December of one year and January of the next year). This discontinuity is allowed to more closely mimic the annual time effects that are conventionally used in the national literature. We also include calendar month dummies (i.e., a dummy for January, February, etc.) to capture seasonal variation.

We estimate the continuation rate at duration k, $c^k(Y_t)$, using individual-level data. The continuation rate is simply one minus the conventional hazard rate for exiting welfare, and thus the estimation problems are equivalent. Let C_{ijt} be an indicator variable equal to one when individual i in county j continues on aid in month t. Consider the model,

$$(6)\ \Pr[C_{ijt}=1\,|\,k]=f(\alpha_c+\beta_c Y_{jt}+\gamma_t+\delta_j+g_c(k_{ijt})),$$

where k_{ijt} is a specification for duration and $g_c(k_{ijt})$ is a flexible specification for the dependence of the continuation probability on k_{ijt}. Again, we choose f to be a linear function due to the size of our data set. Finally, we modify the basic specification by adopting the same specification for the time effects as was adopted for the entry rate model (i.e., a discontinuous piece-wise annual spline and calendar month dummies).

One complication is that we only have data on current welfare receipt status, so we do not know the length of welfare receipt for individuals who are on aid in the first month of our data. To address this form of left censoring, we assume that the probability of continuation becomes constant after $\bar{k}$ periods on aid and then discard the first $\bar{k}$ periods of the data. Therefore, any person continuously on aid from the start of our data to period $\bar{k}+1$ is in the constant part of the hazard, making the left censoring irrelevant. For everyone else, we know the exact duration. The form of the transition matrix in equation (3) includes this assumption.

To motivate our choice for the functional form of $g_c(k_{ijt})$, we present estimates of (6) in which we include twelve lags of the unemployment rate, set $\bar{k}=48$, and specify

$g_c(k_{ijt})$ to be 48 monthly dummies. We graph the regression coefficients of these monthly dummies in Figure 2. It is clear that there is a systematic relationship between the probability of continuing on welfare and duration on welfare. After an initial decline in the probability of continuing, the probability increases relatively smoothly. Given these results, we specify that $g_c(k_{ijt})$ includes dummy variables for the first six months to capture the initial increase and then a quartic in k_{ijt} to capture the subsequent decline. We also initially set $\bar{k}$ to be 24 months and drop the first 24 months of data.

Baseline Results. Table 2 presents the baseline estimates of the effect of economic conditions on entry and continuation (i.e., one minus exit). Our baseline specification sets $\bar{k}$ =24 and includes the current unemployment rate and no lags (columns 1 and 5).

As expected, the entry rate and the continuation rate are counter-cyclical. A higher unemployment rate causes more people to enter welfare and fewer to leave. The contemporaneous effects are clearly statistically significant (t-statistics greater than 4). Furthermore, unlike the common finding for models of the stocks (i.e., the aggregate caseload), there is little evidence of dependence of the transitions on lagged measures of the unemployment rate (models 2 and 6). For the entry rate models, the first and second lags offset each other (the first is positive, the second negative). For the continuation rate, one of the lags is negative (the "wrong sign") and they are both much smaller than the contemporaneous effect (less than a fifth).

We compute the implied effect of the economy on the stock by simulation. Figure 3 presents the results graphically. The heavy line is the path of the actual per capita caseload and the medium line is the path of the simulated caseload based on the actual values of the unemployment rate. As expected, the lines are quite close. The light line gives the path of the caseload implied by the model if the unemployment rate had stayed permanently at its peak of January 1993. In both simulations, the time fixed effects, which capture non-economic factors such as welfare policy, follow their estimated values. By the end of the simulation period, these fixed effects pull down both of the simulated paths, including the one in which the unemployment rate is held at its peak

level. Clearly, non-economic factors have an important role to play in explaining the caseload decline.

Nevertheless, the economy clearly also plays a major role. The implied effect of the economy is given by the difference between the two paths. If the unemployment rate had remained at its peak, the caseload would have continued to rise for approximately another two years before the fixed effects would have pulled it down. Instead, the simulated caseload based on the actual unemployment rate rises only about five percent before peaking in early 1995 and then begins to fall, first slowly and then rapidly.

Table 3 summarizes the effect of the economy on the caseload using our baseline model. The first column of the top panel presents our baseline model ($\bar{k}$=24 with no lags of the unemployment rate). For the actual unemployment rate, the caseload peaks in March of 1995 at 11.1 percent of the population. It then falls to 7.3 percent at the end of our data (December 1998), a total decline of 34.3 percent. The counterfactual simulation (holding the unemployment rate at its maximum of January 1993) implies that over the same period (January 1993 to December 1998) the caseload would have declined to 9.1 percent of the population, representing only an 18.3 percent decline in the caseload. We interpret these results as implying that almost half (46.7 percent) of the caseload decline is due to the economy; i.e. (34.3-18.3)/34.3=46.7.

4. EXTENSIONS OF THE BASIC STOCK-FLOW MODEL

In this section, we consider various extensions to the basic stock-flow model to examine its sensitivity to specification changes. We consider the sensitivity of the results to including different numbers of unemployment rate lags, changing the treatment of initial conditions, allowing the probability of re-entry to differ from the probability of new entry, and allowing the impact of economic conditions to vary by length of time on aid.

Specifying the Number of Unemployment Rate Lags. Klerman and Haider (2001) demonstrate that models of the aggregate caseload are highly sensitive to the number of

lags that were included in the empirical model and provide evidence that this sensitivity arises from the aggregate caseload literature ignoring the underlying stock-flow process. In this section, we present results that include different numbers of unemployment rate lags to examine the sensitivity of the stock-flow model to different numbers of lags.

Table 2 reports regression coefficients for entry (equations 2-4) and continuation (equations 6-8) models with two, five, and eleven lags. In almost all cases, the contemporaneous effect is the largest and has the expected sign. We note that coefficients across lag lengths vary considerably in magnitude, sign, and significance. This pattern is likely due to a combination of measurement error and high serial correlation in the monthly unemployment rate, rather than representing any real economic phenomenon. Using the sum of the coefficients as a proxy for the long-run effect of a permanent change in the unemployment rate, we find that the effects increase as additional lags are added in the entry regressions but decline slightly as additional lags are added in the continuation regressions (see the second to last row of Table 2).

Our primary interest, however, is the implication of these estimates for the aggregate caseload, obtained through simulations. The simulated changes in the caseload suggest that the estimated effect of the economy changes very little when different numbers of unemployment rate lags are included. The results displayed in Table 3 imply that 46.7 percent, 44.7 percent, 45.9 percent, and 46.4 percent of the caseload decline is due to economic conditions using the 0, 2, 5, and 11 lag models, respectively. Thus, the empirical evidence suggests that very few monthly unemployment rate lags are needed to capture the underlying correlations in the data.

The Treatment of Initial Conditions. As described above, we assume that the hazard of leaving welfare is constant after a particular length of time (called $\bar{k}$) to solve an initial conditions problem. The choice of $\bar{k}$ involves balancing two considerations: the larger our choice of $\bar{k}$, the better the model corresponds to the actual pattern in the data (see Figure 2); however, the smaller our choice of $\bar{k}$, the less data we will need to drop to solve the initial conditions problem. The baseline specification sets $\bar{k}=24$. Figure 2

suggests that the true value of $\bar{k}$ is higher than 24 months. The duration hazard coefficients are still increasing at 24 months; however, the coefficients are largely constant after 36 months.

To examine the sensitivity of the simulated effect of the economy to the 24-month assumption in our baseline model, we present simulations in Table 4 for models that are the same as the baseline specification except for setting $\bar{k}$ to be 36 months. Such a specification change requires us to drop an additional twelve months of data as compared to the baseline specification to handle the initial conditions problem. As is clear from the results, the role of the economy becomes slightly more important with just over 50 percent of the decline being attributable to the economy for the $\bar{k}=36$ model. However, these estimates are sufficiently similar that we conclude our baseline assumption of $\bar{k}=24$ is reasonable and that the inclusion of additional lags does not appreciably matter.

Distinguishing New Entry from Re-entry. The basic model implicitly assumes that once a recipient leaves welfare, his or her probability of re-entering welfare is the same as the probability of entry for someone who has never been on welfare. This assumption seems overly restrictive and is different from the specification used by other researchers (e.g., Gittleman 2000). In this sub-section, we allow the probability of re-entry to differ from the probability of new entry and to vary with length of time since last on aid.

A similar initial conditions problem is faced when modeling re-entry as was faced when modeling continuation. In particular, we only observe current welfare status for the period 1987 to 1998. When someone first enters welfare during our sample period, we do not know if they had received welfare prior to 1987. We address this initial conditions problem in a manner analogous to our approach in the continuation regression. Specifically, we assume that the probability of re-entry varies through a specific period, called $\bar{l}$, and thereafter to be identical. By then discarding the first $\bar{l}$ periods of data, all entrants can appropriately be classified as a new entrant (not being on welfare in the last $\bar{l}$ periods) or a re-entry (having been on welfare in the last $\bar{l}$ periods).

We incorporate these extensions by expanding the state-space implicit in equation (3). Specifically, define S_{olt} to be the number of people who have been off aid for l consecutive periods at time t. We allow the re-entry rate (b) to vary with time off aid through $\bar{l}$; after $\bar{l}$, those have been on aid are assumed to be identical to those who have never been on aid. Thus, we re-define (5) to be the hazard of entry for individuals who have not been on aid for the past $\bar{l}$ months and specify a re-entry hazard, similar to equation (6), for all individuals who have been off aid for 1 to $\bar{l}$ months. For the case where $\bar{k} = \bar{l} = 4$, the implied Markov structure can be represented by,

$$(7)\quad \begin{bmatrix} S(r,1,t) \\ S(r,2,t) \\ S(r,3,t) \\ S(r,4,t) \\ S(o,1,t) \\ S(o,2,t) \\ S(o,3,t) \\ S(o,4,t) \\ S(n,t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & b^1 & b^2 & b^3 & b^4 & e \\ c^1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & c^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & c^3 & c^4 & 0 & 0 & 0 & 0 & 0 \\ 1-c^1 & 1-c^2 & 1-c^3 & 1-c^4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-b^1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1-b^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1-b^3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-b^4 & 1-e \end{bmatrix} \begin{bmatrix} S(r,1,t-1) \\ S(r,2,t-1) \\ S(r,3,t-1) \\ S(r,4,t-1) \\ S(o,1,t-1) \\ S(o,2,t-1) \\ S(o,3,t-1) \\ S(o,4,t-1) \\ S(n,t-1) \end{bmatrix}.$$

We estimate the re-entry probability analogously to our approach to estimating continuation. Let R_{ijt} be an indicator variable equal to one for whether individual i in county j re-enters aid month t. Consider the model

$$(8)\quad \Pr[R_{ijt} = 1 \mid l] = f(\alpha_r + \beta_r Y_{ijt} + \gamma_t + \delta_j + g_r(l_{ijt})),$$

where $g_r(l_{ijt})$ is a flexible functional form for duration dependence. To empirically implement the model, we specify the time effects and duration effects as in the continuation rate regressions. Furthermore, we set $\bar{l} = \bar{k} = 24$ and drop the first 24 months of data.

Table 5 presents the entry and re-entry regression results (corresponding to the left panel of Table 2), reclassifying all entries as either being new entries or re-entries. The new entry regressions look qualitatively similar to the baseline entry regressions. The

current unemployment rate effect is about the same. The one and two lag effects are again largely offsetting. The results in columns 4 through 6 show that re-entry is also affected by the unemployment rate.

Our primary interest is the implied overall change in the role of the economy. Table 6 gives the results of the simulation of the stock when separating new entry from re-entry, similar to the results in Table 3 where new entry and re-entry were not treated separately. As is readily apparent, the results are very similar. For example, the results in Table 3 and Table 6 imply similar roles for the economy across all four models (42.2 to 47.5 percent in the models that account for reentry, versus 44.7 to 46.7 percent in the models that do not), although the exact ranking across the models has changed. Thus, our results suggest that the processes of entry and re-entry are different but separating the two is not empirically important from the perspective of understanding the impact of economic conditions on the welfare caseload.

Allowing Differential Impacts of the Unemployment Rate by Duration. The baseline model implicitly assumes that the impact of the economy is invariant with respect to how long individuals have been on aid, i.e., that there is no interaction between duration and the economy. Given the strong duration dependence found in Figure 2, it is plausible that the effect of the economy (as proxied by the unemployment rate) also varies with duration. For example, long-term welfare users might be less able to find employment during economic expansions because of human capital depreciation from being out of the labor market or due to stigma associated with being on welfare. We explore this possibility in this sub-section.

Based on the results in the previous sub-section, we define our base model to include only the contemporaneous unemployment rate. We then allow the impact of the unemployment rate to vary by duration with two different specifications. In the first specification, we include interactions with three duration dummies (on aid for 0-11 months, 12-23 months, and 24+ months). In the second specification, we include an interaction of the unemployment rate with a continuous measure of time on aid (called k).

In the specification with dummy variables (column 2 of Table 7), the coefficients in the continuation models are essentially identical, with coefficients of 0.114, 0.103, and 0.112 for the 0-11 month, 12-23 month, and 24+ month interactions, respectively (each with standard errors of approximately 0.006). The second specification (in column 2) in which the effect of the economy is allowed to vary linearly with duration gives equivalent results. The point estimate on the interaction term is zero to three decimal places.

We simulate the implications for these flow relationships to changes in the caseload stock and present the results in Table 8. Given the flow regressions, it is not surprising that the implied effects of the economy do not change much as compared to the baseline results in Table 3. In particular, the role for the economy is still approximately 40 percent (see Table 8). Therefore, these results suggest that the base-case assumption that the relationship between continuation and the unemployment rate is additive and invariant across durations is consistent with the empirical evidence.

5. ALTERNATIVE MEASURES OF ECONOMIC CONDITIONS

Following the previous literature, our baseline model uses the unemployment rate as its proxy for economic conditions. However, there are at least three reasons why the unemployment rate might be an incomplete proxy. First, the unemployment rate is likely to be measured with error, particularly given that we use monthly county-level estimates. The county-level measures that we use are produced as part of the joint federal-state Local Area Unemployment Statistics program, administered by the Bureau of Labor Statistics. These unemployment rates are constructed through a synthetic estimation process that relies on state-level estimates derived from the Current Population Survey and is then extended to smaller areas using Current Employment Statistics and state unemployment insurance data. It seems likely that this process yields estimates with considerable measurement error.[11] To the extent that the measurement error is serially

[11] It should be noted that the state level unemployment estimates are also presumably estimated with measurement error, particularly for the states with a small population. The state unemployment estimates are largely based on the Current Population Survey (CPS), and the CPS sample is small for the small states.

uncorrelated, the coefficient on the unemployment rate will tend to suffer from attenuation bias, i.e., point estimates that are too small in absolute value. Including multiple lags of the unemployment rate (as we did above) will partially correct our simulations for such attenuation bias.

Second, even if perfectly measured, the unemployment rate might be the wrong measure of economic opportunity for the welfare population. As is well known, the Bureau of Labor Statistics definition of unemployed requires that an individual be actively seeking employment. If individuals who go on welfare do not seek employment, then counties with high welfare rolls might have a low unemployment rate even though individuals initially go onto welfare because of the lack of economic opportunity.

Third, and more generally, it seems likely that the economy is not a one-dimensional concept. For example, the labor market opportunities might change differentially for a low-skilled person versus a high-skilled person, and a uni-dimensional measure such as the unemployment rate could not capture differential movements. Moreover, it is not just the probability of working that is important, but also the wages that would be received when working. The idea of treating the economy as a multi-dimensional concept and using multiple proxies is consistent with other work in the regional economics literature (Bartik, 1991; Batik, 1996; Blanchard and Katz, 1992) and in labor economics (Davis and Haltiwanger, 1992; Katz and Murphy, 1992).

The argument for a multi-dimensional conceptualization of the economy seems particularly strong for the population most at risk of entering welfare. The high-risk population tends to be young, single women with little experience and few skills. For such individuals, labor market outcomes at the bottom of the skill distribution seem most relevant, and tightness or slack at the top or even middle of the labor market is likely to be less important. Of course, we expect that outcomes for the bottom of the labor market will be positively correlated with outcomes for the labor market as a whole. It is, however, far from clear how strong that correlation will be. The literature on increasing earnings inequality (e.g., Katz and Murphy, 1992; Juhn, 1992; Juhn, Murphy, and Pierce, 1993) suggests that labor market outcomes have diverged widely at the top and bottom of

the distribution. Consistent with this perspective, some of the papers in the regional economics literature find differential effects of various proxies across segments of the labor market (e.g., Bartik, 1991; 1996; Blanchard and Katz, 1992). Thus, we would like to have a proxy for the economy that is more targeted at the bottom of the labor market.

Three other papers of which we are aware have made similar arguments and explore an expanded set of proxies for economic conditions. Each of them finds an important role for additional measures of the economy. Bartik and Eberts (1999) estimate models similar to the conventional caseload literature (using national data and including state and year fixed effects) and find strong effects of the employment growth rate, even when the unemployment rate is also included. They find some evidence to support the inclusion of current or lagged values of the employment rate for high school graduates, the wage premium implied by a state's industry composition, and the employment rate implied by a state's industry composition. Ribar (2000) constructs a county-level index of labor market opportunity for low-skilled individuals and finds that the index predicts caseload transitions significantly better than the local area unemployment rate.

Most relevant to our work is Hoynes (2000), who also uses California MEDS data. She explores the effect of economic conditions on exit from and re-entry onto welfare. She does not examine the impact of economic conditions on entry, and thus cannot explore the effect of the economy on the total caseload. Hoynes (2000) finds that average earnings overall, in services, and in retail trade have a significant effect on continuation and re-entry, over and above the unemployment rate. In fact, when she includes county and time fixed effects (as we do), inclusion of the earnings measures make the effect of the unemployment rate much smaller, statistically indistinguishable from zero. She also considers models with employment (total and by sector) and the employment-to-population ratio (total and by sector). Overall she concludes "models that control for labor market conditions using employment-based measures perform better than unemployment rates."

Four Measures of Economic Conditions. To explore the sensitivity of our results, we consider four additional measures: the employment rate, the employment growth rate,

earnings per worker in the retail trade sector, and the growth rate of earnings per worker in the retail trade sector.

These additional measures are intended to remedy some of the problems with the unemployment rate previously discussed. The employment rate is similar to the unemployment rate in that it is intended to measure the broad, economic activity in the county; however, the employment rate should be measured with less error because it is based on administrative data compiled as part of the unemployment insurance system. Moreover, it is not sensitive to a discouraged worker effect. The employment growth rate is intended to be a measure of the slack in the labor market, rather than a measure of the general economic activity (see Bartik and Eberts, 1999).

The measures of earnings in retail trade are intended to provide additional information about the opportunities in the labor market for welfare recipients. Earnings per worker in the retail trade sector are intended to provide a more specific measure of the opportunities available to this group. We also include the growth rate in earnings as another measure of slack in the labor market.

Table 9 presents a correlation matrix for the various economic condition measures. Given that many of the underlying concepts are conceptually closely related, the actual correlations are small. The largest correlation (in absolute value) is between the earnings per worker and the growth rate in earnings per worker (-0.74), followed by the correlation between the unemployment rate and the growth rate in earnings per worker (0.51). Notably, the correlation between the unemployment rate and the employment growth rate is very small (-0.05). Thus, the simple correlations suggest that the various measures are somewhat distinct.

Results with Alternative Measures. Turning to the effects on the transition rates, Table 10 presents results for entry rate regressions and Table 11 presents results for continuation rate regressions. For both sets of regressions, each of the additional measures of economic conditions enters with the expected sign and is statistically significant. The only exception to this pattern is that the growth rate in earnings per

worker is not significant in the model when all measures are included simultaneously in the entry regressions (column 4 in Table 10).

Again, our main interest is in the total impact of the economy on the per capita caseload. In Table 12, we present simulations based on the flow estimates of Tables 10 and 11. Unlike our other sensitivity analyses, Table 12 suggests that the simulated effect of the economy is very sensitive to changes in how the economy is measured. In the baseline specification, the economy explains 46.7 percent (0 lag model and $\bar{k} = 24$ from Table 3). Including the employment rate and the employment growth rate has little effect on the total estimate, with the estimated role for the economy changing marginally to 46.2 percent. Strikingly, including per worker earnings in retail trade causes a sharp increase in the effect of the economy. For example, in the model that includes earnings and the earnings growth rate, we compute that fraction attributable to the economy approximately doubles to 93.7 percent (see model 2). The results remain at a similar level with various other combinations of economic conditions, including all of the measures (model 4). Therefore, we conclude that approximately ninety percent of the caseload decline in California from 1995 to 1998 is due to economic conditions.

Substantively, two aspects of these results deserve comment. First, these results suggest that the wage level provides important additional information about economic conditions not captured by the unemployment rate. Second, welfare reform was not passed in California until 1998 and then was only slowly implemented over the year; thus, it is unlikely that these changes in the unemployment rate or earnings could have been caused by policy changes. Moreover, it is likely that the role for the economy in other states would be smaller because California adopted its policy changes relatively late and the changes were less dramatic than those in other states.

6. DISCUSSION AND CONCLUSION

Klerman and Haider (2001) propose a stock-flow model of the welfare caseload and then estimate the model using California data. This paper has explored the sensitivity of the stock-flow approach to two sets of specification choices. First, the paper explored the

sensitivity of the empirical model to several structural changes, including the treatment of initial conditions, distinguishing entry from re-entry, and allowing economic effects on continuation to vary with duration. None of the substantive conclusions of the basic model change. In particular, regardless of the specification changes, the economy explains approximately half of California's 31 percent caseload decline from its peak in March 1995 to the end of our data in December 1998.

In contrast to this null result, varying the measure of economic conditions has a large effect on the substantive results. Following the majority of previous research, our baseline model uses the conventional unemployment rate as a proxy for economic conditions. Consistent with other studies that have also augmented the measure of economic conditions, we find that the unemployment rate alone is not sufficient to capture the effect of economic conditions on the welfare caseload. Furthermore, including per worker earnings as a measure of economic conditions doubles the estimated role for the economy. In the models that include earnings in retail trade, the economy explains more than 90 percent of the caseload decline. We conclude that a dynamic model with a rich parameterization for economic conditions can explain the overwhelming share of the caseload decline in California over this period. Furthermore, this estimated effect is much larger than that obtained either from aggregate regressions or from our stock-flow approach when only the unemployment rate is used as a proxy for economic conditions.

Combined with the comparison between conventional stock models and the stock-flow models reported in Klerman and Haider (2001), our results have important implications for understanding the causes of the decline in the U.S. welfare caseload in the late-1990s. Klerman and Haider (2001) find that, when using the unemployment rate alone, the stock-flow approach yields estimated effects of the economy larger than those implied by the conventional stock models (e.g., CEA 1997). The results presented here suggest that the basic stock-flow analysis is robust to many different changes in the structure of the model, but expanding the measures for the economy substantially increases the estimated role of the economy in the caseload decline.

These results suggest several broader conclusions. First, because we find an even larger role for the economy, it is more likely that there will be substantial increases in the caseload with the next recession. Second, even though we do not examine the impact of policy directly, our results have important implications for understanding the potential role for policy. As the estimated role of the economy expands, the potential role for welfare reform contracts. We note, however, that the specific percentage estimates of the decline due to the economy presented here should not necessarily be applied directly to other states. In particularly, the process of welfare reform began later and was less drastic in California than in the rest of the United States and our data only cover the period through the end of 1998, so it is possible that policy changes had a greater impact in other states and in California for a later period.

Finally, our results suggest that it is possible and useful to model the underlying dynamics directly and to derive the implications for the stock by simulation. Arguments similar to those made here also apply to the impact of policy on the welfare caseload. Moreover, it is likely that similar dynamics are relevant to many other government assistance programs, such as Food Stamps and Medicaid (Schoeni, 2001). Finally, many other processes that are inherently persistent should be expected to evolve similarly, such as occupational structures and prison populations.

REFERENCES

Bane, Mary Jo and David Ellwood (1986). "Slipping Into and Out of Poverty: The Dynamics of Spells." *Journal of Human Resources*; 21(1), Winter 1996, 1-23.

Bartik, Timothy J. (1991). *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, Michigan: W.E. Upjohn Institute for Employment Research.

Bartik, Timothy J. (1996). "The Distributional Effects of Local Labor Demand and Industrial Mix: Estimates Using Individual Panel Data." *Journal of Urban Economics*, 40: 150-178.

Bartik, Timothy J., and Randall W. Eberts (1999). "Examining the Effect of Industry Trends and Structure on Welfare Caseloads," mimeograph, November, W.E. Upjohn Institute for Employment Research.

Blanchard, O.J. and L.F. Katz (1992). "Regional Evolutions," *Brookings Papers on Economic Activities*, 1: 1-75.

Blank, R. and P. Ruggles (1996). "When Do Women Use AFDC and Food Stamps? The Dynamics of Eligibility vs. Participation," *Journal of Human Resources,* 31(1):57-89.

Council of Economic Advisers (1997). "Technical Report: Explaining the Decline in Welfare Receipt, 1993-96." A Report by the Council of Economic Advisors, May.

Council of Economic Advisers (1999). "Technical Report: The Effects of Welfare Policy and the Economic Expansion on Welfare Caseloads: An Update." A Report by the Council of Economic Advisors, August.

Currie, Janet and Jeff Grogger (2000). "Explaining Recent Declines in the Food Stamp Program Participation." September manuscript.

Davis, Steven J. and John C. Haltiwanger (1992). "Gross Job Creation, Gross Job Destruction, and Employment Reallocation." *Quarterly Journal of Economics*, 107:819-863.

Figlio, David and James Ziliak (1999). "Welfare Reform, the Business Cycle, and the Decline in AFDC Caseloads." In *Economic Conditions and Welfare Reform* (ed. Sheldon Danziger), Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Gittleman, Maury (2000). "Declining Caseloads: What Do the Dynamics of Welfare Participation Reveal?" Unpublished manuscript.

Haider, Steven, Jacob Alex Klerman, Jan Hanley, Laurie McDonald, Elizabeth Roth, Liisa Hiatt, and Marika Suttorp (1999). "Welfare Reform in California: Design of the Impact Analysis, Preliminary Investiations of Caseload Data." Draft of DRR(L)-2077/1-Cdss, May.

Haider, Steven and Jacob Alex Klerman (2001). "Dynamic Properties of the Welfare Caseload." Unpublished manuscript.

Heckman, James and James Walker (1990). "The Relationship between Wages and Income and the Timing and Spacing of Births: Evidence from Swedish Longitudinal Data." *Econometrica*, 58(6) 1411-1441.

Hoynes, Hilary Wiliamson (2000). "Local Labor Markets and Welfare Spells: Do Demand Conditions Matter?" *Review of Economics and Statistics*, 82(3) 351-368.

Hoynes, Hilary Williamson and Thomas MaCurdy (1994). "Has the Decline in Benefits Shortened Welfare Spells?" *American Economic Review*, 84(2) 43-48.

Hutchens, Robert (1981). "Entry and Exit Transitions in a Government Transfer Program: The Case of Aid to Families with Dependent Children." *Journal of Human Resources*, 16(2).

Katz, Lawrence F. and Kevin M. Murphy (1992). "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *Quarterly Journal of Economics*, 107(1) 35-78.

Juhn, Chinhui (1992). "Decline of Male Labor Market Participation: The Role of Declining Market Opportunities." *Quarterly Journal of Economics*, 107(1) 79-121.

Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce (1993). "Wage Inequality and the Rise in Returns to Skill." *The Journal of Political Economy*, 101(3), 410-442.

Levine, Phillip and Diane Whitmore (1998). "The Impact of Welfare Reform on the AFDC Caseload" *National Tax Association Proceedings*. Washington, DC: Nation Tax Association, 24-33.

Maddala, G. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

Moffitt, Robert (1992). "Incentive Effects of the U.S. Welfare System: A Review." *Journal of Economic Literature*, 30, 1-61.

Moffitt, Robert (1999). "The Effect of Pre-PRWORA Waivers on AFDC Caseloads and Female Earnings, Income, and Labor Force Behavior." May manuscript.

Moffitt, Robert and Michael Rendall (1995). "Cohort Trends in the Lifetime Distribution of Female Family Headship in the United States, 1968-1985." *Demography*, 32(3) (Aug.), 407-424.

Ribar, David (2000). "Transitions from Welfare and the Employment Prospects of Low-skill Workers." March manuscript.

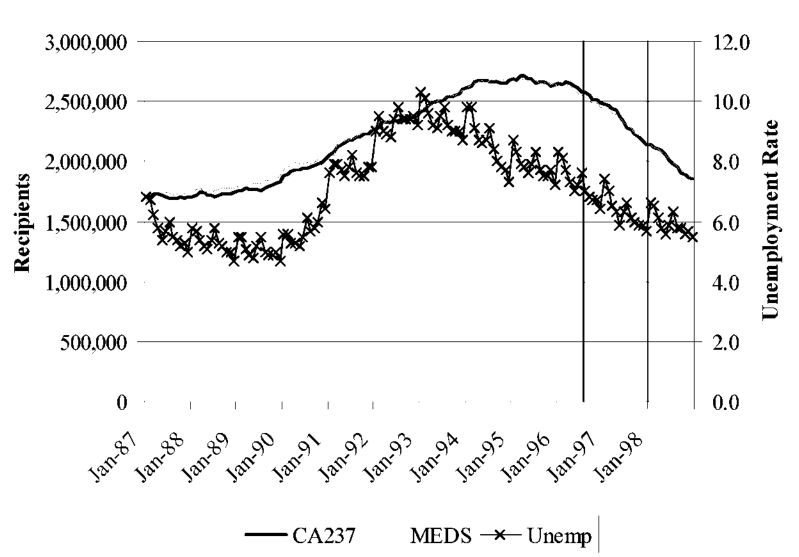
Schoeni, Robert F. (2001). "Comment on: Currie and Grogger, Explaining Recent Declines in Food Stamp Program Participation," in *Brookings and Wharton Papers on Urban Affairs*, William G. Gale and Janet Rothenberg Pack, ed. Brookings Institution Press: Washington, DC.

Schoeni, Robert and Rebecca Blank (2001). "What Has Welfare Reform Accomplished? Impacts on Welfare Participation, Employment, Income, Poverty, and Family Structure." September manuscript.

Wallace, Geoffrey and Rebecca Blank (1999). "What Goes Up Must Come Down?" In *Economic Conditions and Welfare Reform* (ed. Sheldon Danziger), Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

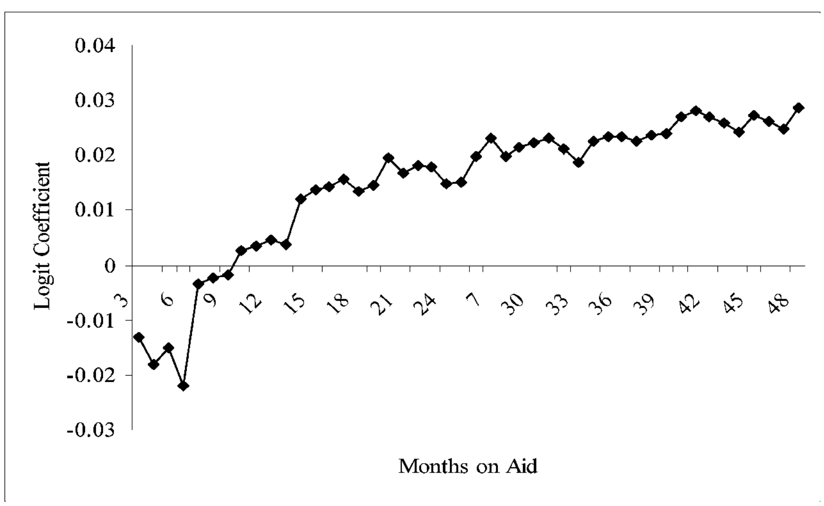
Ziliak, James, David Figlio, Elizabeth Davis, Laura Connolly (1997). "Accounting for the Decline in AFDC Caseloads: Welfare Reform or Economic Growth?" Institute for Research on Poverty Discussion Paper No. 1151-97.

Figure 1: Welfare Recipients and Unemployment Rate in California



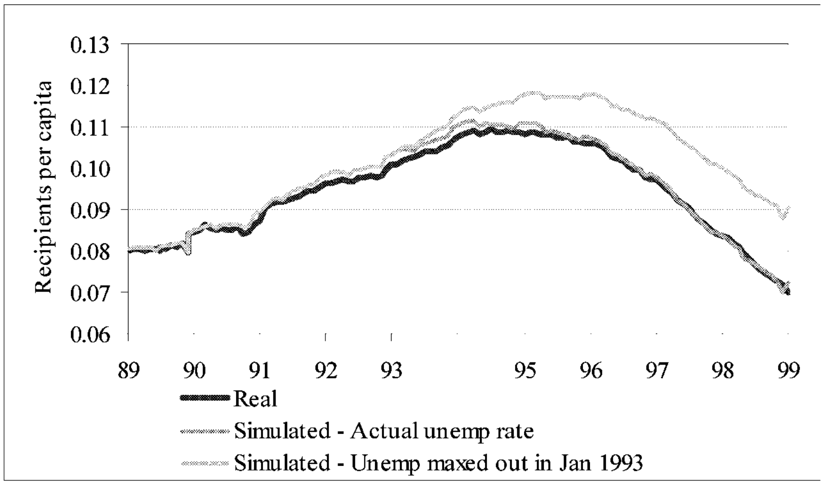
Note: Authors' computations from the MEDS, CA237, and the California unemployment rate. The CA237 is the official California welfare caseload. The MEDS represents an estimate of the caseload based on the 3% sample we analyze in this paper. The first vertical line represents in the passage of the federal welfare reform legislation (PRWORA, August 1996) and the second vertical line represents the implementation of the California welfare reform (CalWORKs, January 1998).

Figure 2: The Variation in Continuation by Duration on Aid



Note: Authors' computations from the MEDS. This figure graphs the logit coefficients for 47 dummy variables for time on aid from the continuation regression; *kbar* is set to 48. The model contains county fixed effects and a flexible spline in time. See the text for further details. Standard errors are in parentheses.

Figure 3: The Simulated Welfare Caseload in California--Holding the Unemployment Rate Constant in January 1993



Note: Authors' computations from the MEDS. The three lines correspond to the real number of recipients per capita over time, and simulations based on the actual unemployment rate (mean of 0-2 lags in the unemployment rate) and based on holding the unemployment rate held constant after January 1993 (mean of 0-2 lags in the unemployment rate).

Table 1: Short Term Spell Durations, 1989-1997

Spell Start Period	*Spells*	*Average Monthly Entry Rate*	*Average Monthly Continuation Rate for Spells that Lasted:* 2-5 mos.	6-11 mos.	12-17 mos.	18+ mos.
1/89-12/90	70,721	0.0032	0.938	0.943	0.964	0.979
1/91-12/92	79,620	0.0037	0.942	0.950	0.968	0.982
1/93-12/94	80,863	0.0037	0.946	0.951	0.968	0.982
1/95-12/96	72,234	0.0031	0.942	0.949	0.962	0.977
1/97-12/97	29,862	0.0024	0.933	0.942	X	X

Source: Authors' tabulations from the MEDS.
Note: An "X" indicates that the probability could not be calculated because of right-censoring.

Table 2: Probability of Entry and Continuation Regressions

	Probability of Entry (Grouped Logit)				*Probability of Continuation (OLS: coeffs & s.e. * 100)*			
Regressors	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Unemp. Rate	0.034	0.029	0.035	0.022	0.111	0.126	0.120	0.111
	(0.003)	(0.006)	(0.006)	(0.007)	(0.005)	(0.009)	(0.010)	(0.011)
UR-1st lag		0.019	0.024	0.030		-0.024	-0.040	-0.046
		(0.008)	(0.008)	(0.009)		(0.013)	(0.013)	(0.014)
UR-2nd lag		-0.019	-0.038	-0.038		0.007	0.074	0.089
		(0.006)	(0.009)	(0.009)		(0.009)	(0.014)	(0.014)
UR-3rd lag			0.010	0.013			-0.100	-0.095
			(0.009)	(0.009)			(0.014)	(0.014)
UR-4th lag			0.005	0.016			0.075	0.067
			(0.008)	(0.009)			(0.013)	(0.013)
UR-5th lag			0.006	-0.013			-0.036	-0.018
			(0.006)	(0.009)			(0.009)	(0.014)
UR-6th lag				0.014				-0.026
				(0.009)				(0.014)
UR-7th lag				0.001				-0.029
				(0.009)				(0.014)
UR-8th lag				-0.007				0.063
				(0.009)				(0.014)
UR-9th lag				0.008				-0.012
				(0.008)				(0.014)
UR-10th lag				0.012				-0.005
				(0.008)				(0.013)
UR-11th lag				0.003				0.0009
				(0.007)				(0.0101)
Sum of coeffs.	0.034	0.029	0.042	0.061	0.111	0.109	0.092	0.100
R-squared	0.8220	0.8223	0.8228	0.8238	0.0096	0.0096	0.0096	0.0096
Observations	6,480	6,480	6,480	6,480	9,080,952	9,070,952	9,070,952	9,070,952

Source: Authors' computations from the MEDS.
Note: We use monthly data for the period January 1989 to December 1998 for each set of models, with the lags referring to monthly lags; *kbar* is set to 24 for the continuation models. All models contain county fixed effects and a flexible spline in time. See the text for further details. Standard errors are in parentheses.

Table 3: Basic Stock-Flow Simulations of the Role of the Economy, $\bar{k}$ =24

Simulations	*0 Lag Model*	*2 Lag Model*	*5 Lag Model*	*11 Lag Model*
(1) Actual unemployment rate				
Simulated March 1995 level	0.111	0.111	0.111	0.111
Simulated December 1998 level	0.073	0.073	0.073	0.073
Simulated percent decline	-34.3%	-34.3%	-34.3%	-34.2%
(2) Unemp. rate remains constant after 1/93				
Simulated December 1998 level	0.091	0.090	0.090	0.091
Simulated percent decline	-18.3%	-19.0%	-18.5%	-18.4%
(3) Decline attributable to economic conditions	46.7%	44.7%	45.9%	46.4%

Source: Authors' computations from the MEDS.
Note: Simulations are based on the stock-flow model. Calculations are based on monthly data for the period January 1989 to December 1998. See the text further details.

Table 4: Basic Stock-Flow Simulations of the Role of the Economy, $\bar{k}$ =36

Simulations	*0 Lag Model*	*2 Lag Model*	*5 Lag Model*	*11 Lag Model*
(1) Actual unemployment rate				
Simulated March 1995 level	0.114	0.114	0.114	0.114
Simulated December 1998 level	0.075	0.075	0.075	0.075
Simulated percent decline	-34.1%	-34.1%	-34.1%	-34.0%
(2) Unemp. rate remains constant after 1/93				
Simulated December 1998 level	0.095	0.095	0.096	0.098
Simulated percent decline	-16.7%	-17.0%	-15.3%	-14.1%
(3) Decline attributable to economic conditions	51.1%	50.2%	55.0%	58.5%

Source: Authors' computations from the MEDS.
Note: Simulations are based on the stock-flow model. Calculations are based on monthly data for the period January 1990 to December 1998. See the text further details.

Table 5: Probability of New Entry versus Re-Entry

	Probability of New Entry (Grouped Logit)				*Probability of Re-Entry (OLS: coeffs & s.e. * 100)*			
Regressors	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Unemp. Rate	0.036	0.037	0.048	0.025	0.026	0.088	0.008	-0.045
	(0.004)	(0.007)	(0.007)	(0.008)	(0.005)	(0.009)	(0.010)	(0.011)
UR-1st lag		-0.005	-0.001	0.006		-0.074	-0.090	0.103
		(0.010)	(0.010)	(0.010)		(0.013)	(0.014)	(0.014)
UR-2nd lag		0.007	-0.005	-0.002		-0.008	0.041	-0.033
		(0.007)	(0.010)	(0.010)		(0.009)	(0.014)	(0.014)
UR-3rd lag			-0.004	0.001			0.044	-0.031
			(0.010)	(0.010)			(0.014)	(0.014)
UR-4th lag			-0.002	0.009			-0.157	0.140
			(0.010)	(0.010)			(0.013)	(0.014)
UR-5th lag			0.029	0.009			0.063	-0.073
			(0.007)	(0.010)			(0.009)	(0.014)
UR-6th lag				0.018				-0.030
				(0.011)				(0.014)
UR-7th lag				-0.008				0.033
				(0.010)				(0.014)
UR-8th lag				-0.005				0.058
				(0.010)				(0.014)
UR-9th lag				0.025				-0.070
				(0.010)				(0.014)
UR-10th lag				-0.010				0.093
				(0.010)				(0.013)
UR-11th lag				0.023				-0.112
				(0.008)				(0.010)
Sum of coeffs.	0.036	0.038	0.066	0.090	0.026	0.006	-0.024	0.034
R-squared	0.7799	0.7800	0.7814	0.7831	0.0228	0.0228	0.0228	0.0228
Observations	6,480	6,480	6,480	6,480	6,275,041	6,275,041	6,275,041	6,275,041

Source: Authors' computations from the MEDS.
Note: We use monthly data for the period January 1989 to December 1998 for each set of models, with *kbar* set to 24 for the re-entry models. All models contain county fixed effects and a flexible spline in time. See the text for further details. Standard errors are in parentheses.

Table 6: Simulations that Distinguish between New Entry and Re-entry

Simulations	*0 Lag Model*	*2 Lag Model*	*5 Lag Model*	*11 Lag Model*
(1) Actual unemployment rate				
Simulated March 1995 level	0.112	0.112	0.112	0.112
Simulated December 1998 level	0.074	0.074	0.075	0.075
Simulated percent decline	-33.3%	-33.3%	-33.3%	-33.2%
(2) Unemp. rate remains constant after 1/93				
Simulated December 1998 level	0.090	0.091	0.092	0.092
Simulated percent decline	-19.3%	-18.1%	-17.5%	-17.8%
(3) Decline attributable to economic conditions	42.2%	45.7%	47.5%	46.3%

Source: Authors' computations from the MEDS.
Note: We use monthly data for the period January 1989 to December 1998 for each set of models, with *kbar* set to 24 for the continuation and re-entry models. The calculations are based on the flow results in Tables 2 and 5.

Table 7: Continuation Regressions Interacting Unemployment with Duration

	*Probability of Contiuation (OLS: coeffs & s.e. * 100)*	
Regressors	*1*	*2*
Unemp0		0.117
		(0.006)
Unemp0*dur0-11	0.114	
	(0.006)	
Unemp0*dur11-23	0.103	
	(0.006)	
Unemp0*dur24+	0.112	
	(0.005)	
Unemp0*t		$-3*10^{-4}$
		$(2*10^{-4})$
R-squared	0.0096	0.0096
Observations	9,070,952	9,070,952

Source: Authors' computations from the MEDS.
Note: We use monthly data for the period January 1989 to December 1998 for each set of models; *kbar* is set to 24 for the continuation models. All models contain county fixed effects and a flexible spline in time. See the text for further details. Standard errors are in parentheses.

Table 8: Simulations that Allow for Duration Interactions

Simulations	*Baseline*	*1*	*2*
(1) Actual unemployment rate			
Simulated March 1995 level	0.111	0.111	0.111
Simulated December 1998 level	0.073	0.073	0.073
Simulated percent decline	-34.3%	-34.3%	-34.3%
(2) Unemp. rate remains constant after 1/93			
Simulated December 1998 level	0.091	0.088	0.088
Simulated percent decline	-18.3%	-20.4%	-20.5%
(3) Decline attributable to economic conditions	46.7%	40.5%	40.3%

Source: Authors' computations from the MEDS.
Note: We use monthly data for the period January 1989 to December 1998 for each set of models, with *kbar* set to 24 for the continuation models. The calculations are based on the flow results in Tables 2 and 7.

Table 9: Correlations Among Various Measures of Economic Conditions

	Unemploy-ment Rate	Employment rate	Employment growth rate	Per-worker earnings	Per worker earnings growth
Unemployment Rate	1				
Employment rate	-0.397 <.0001	1			
Employment rate growth rate	-0.052 <.0001	0.044 0.0004	1		
Per-worker earnings (retail)	-0.251 <.0001	0.114 <.0001	0.033 0.008	1	
Per worker earnings (retail) growth rate	0.050 <.0001	-0.265 <.0001	-0.009 0.478	-0.741 <.0001	1
Observations	6480	6480	6480	6480	6480

Source: Authors' computations from the unemployment rate.
Note: We use monthly data for the period January 1989 to December 1998. The calculations are weighted by county population.

Table 10: Entry Regressions using Other Measures of Economic Conditions

	Probability of Entry (Grouped Logit)			
Regressors	*1*	*2*	*3*	*4*
Unemployment rate	0.018 (0.003)	0.030 (0.003)	0.031 (0.003)	0.015 (0.004)
Employment rate	-0.018 (0.002)			-0.016 (0.002)
Employment growth rate	0.005 (0.002)		0.002 (0.002)	0.005 (0.002)
Per-worker earnings (retail)		-2.649 (0.715)	-1.220 (0.150)	-1.608 (0.719)
Per worker earnings (retail) growth rate		-0.144 (0.070)		-0.049 (0.070)
R-squared	0.8249	0.8239	0.8238	0.8264
Observations	6480	6480	6480	6480

Source: Authors' computations from the MEDS.
Note: We use monthly data for the period January 1989 to December 1998. All models contain county fixed effects and a flexible spline in time. See the text for further details. Standard errors are in parentheses.

Table 11: Continuation Regressions using Alternative Measures of Economic Conditions

	*(OLS: both coeffs & s.e. * 100)*			
Regressors	*1*	*2*	*3*	*4*
Unemployment rate	0.116	0.104	0.098	0.115
	(0.006)	(0.005)	(0.005)	(0.006)
Employment rate	0.009			0.015
	(0.003)			(0.003)
Employment growth	-0.018		-0.017	-0.019
rate	(0.004)		(0.004)	(0.004)
Per-worker earnings		-15.291	-3.530	-16.025
(retail)		(1.430)	(0.308)	(1.444)
Per worker earnings		-1.155		-1.218
(retail) growth rate		(0.137)		(0.138)
R-squared	0.0096	0.0096	0.0096	0.0096
Observations	9,070,952	9,070,952	9,070,952	9,070,952

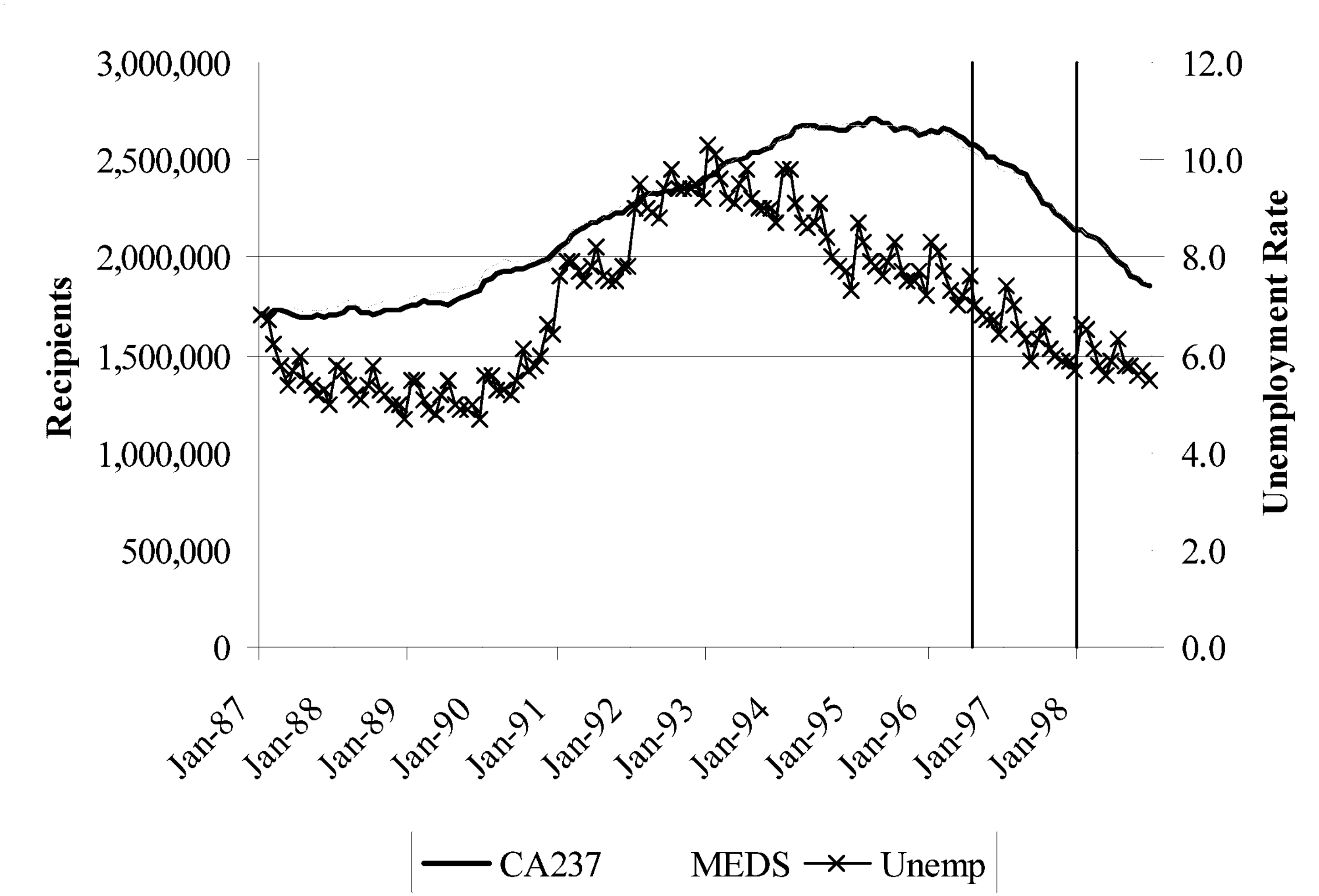
Source: Authors' computations from the MEDS.
Note: We use monthly data for the period January 1989 to December 1998 for each set of models; *kbar* is set to 24 for the continuation models. All models contain county fixed effects and a flexible spline in time. See the text for further details. Standard errors are in parentheses.

Table 12: Simulations with Alternative Measures of Economic Conditions

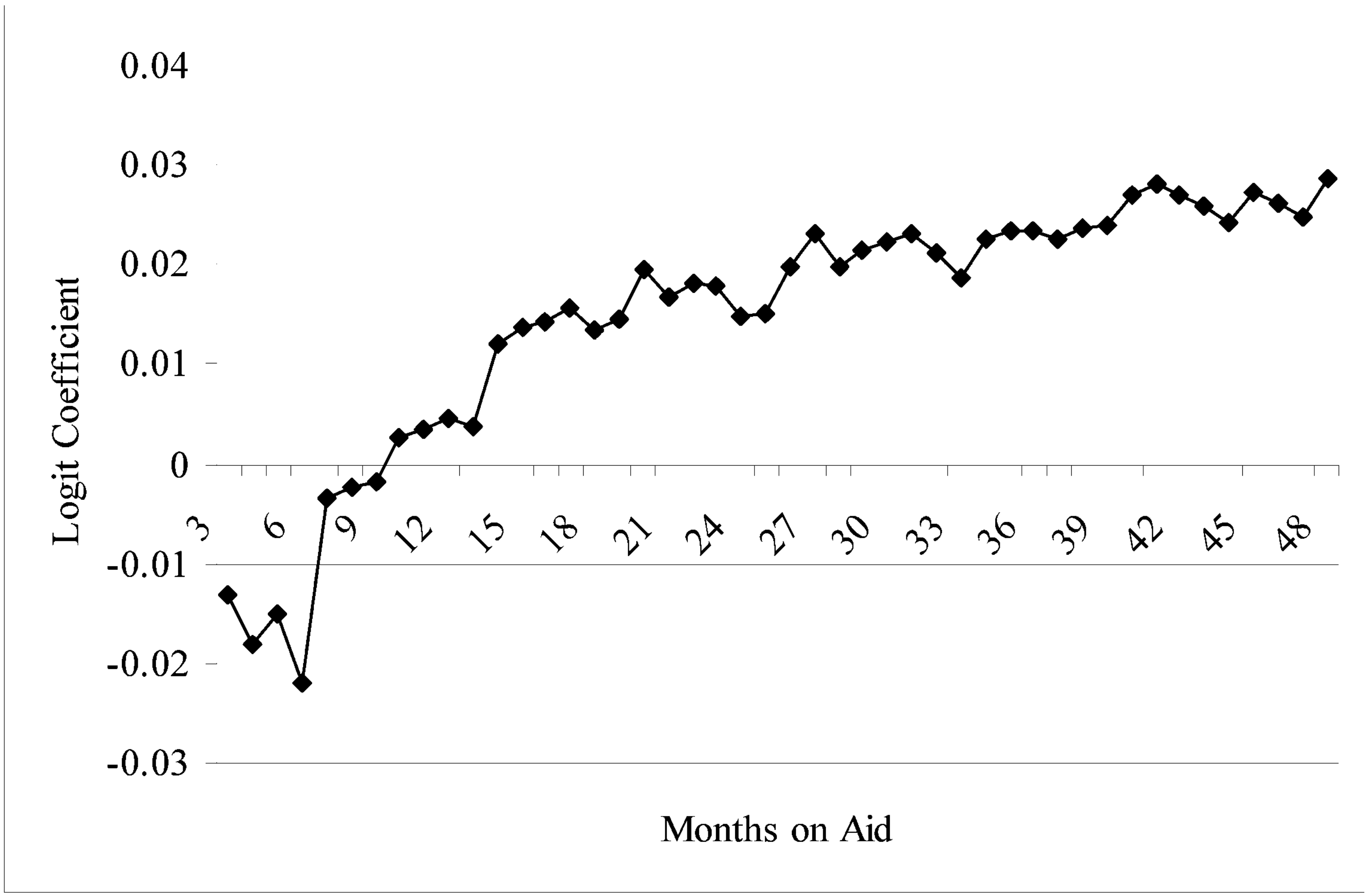
Simulations	*Baseline*	*1*	*2*	*3*	*4*
(1) Actual unemployment rate					
Simulated March 1995 level	0.111	0.111	0.111	0.111	0.111
Simulated December 1998 level	0.073	0.073	0.073	0.073	0.073
Simulated percent decline	-34.3%	-34.1%	-34.0%	-34.1%	-33.8%
(2) Unemp. rate remains constant after unemployment rate peak (1/93)					
Simulated December 1998 level	0.091	0.090	0.108	0.108	0.108
Simulated percent decline	-18.3%	-18.4%	-2.1%	-3.0%	-2.1%
(3) Decline Attributable to Economic Conditions	46.7%	46.2%	93.7%	91.3%	93.7%

Source: Authors' computations from the MEDS.
Note: We use monthly data for the period January 1989 to December 1998 for each set of models, with *kbar* set to 24 for the continuation models. The baseline model corresponds to the results from the 0 lag model in Table 3. Models 1 through 4 correspond to the flow models 1 through 4 in Tables 10 and 11, respectively.

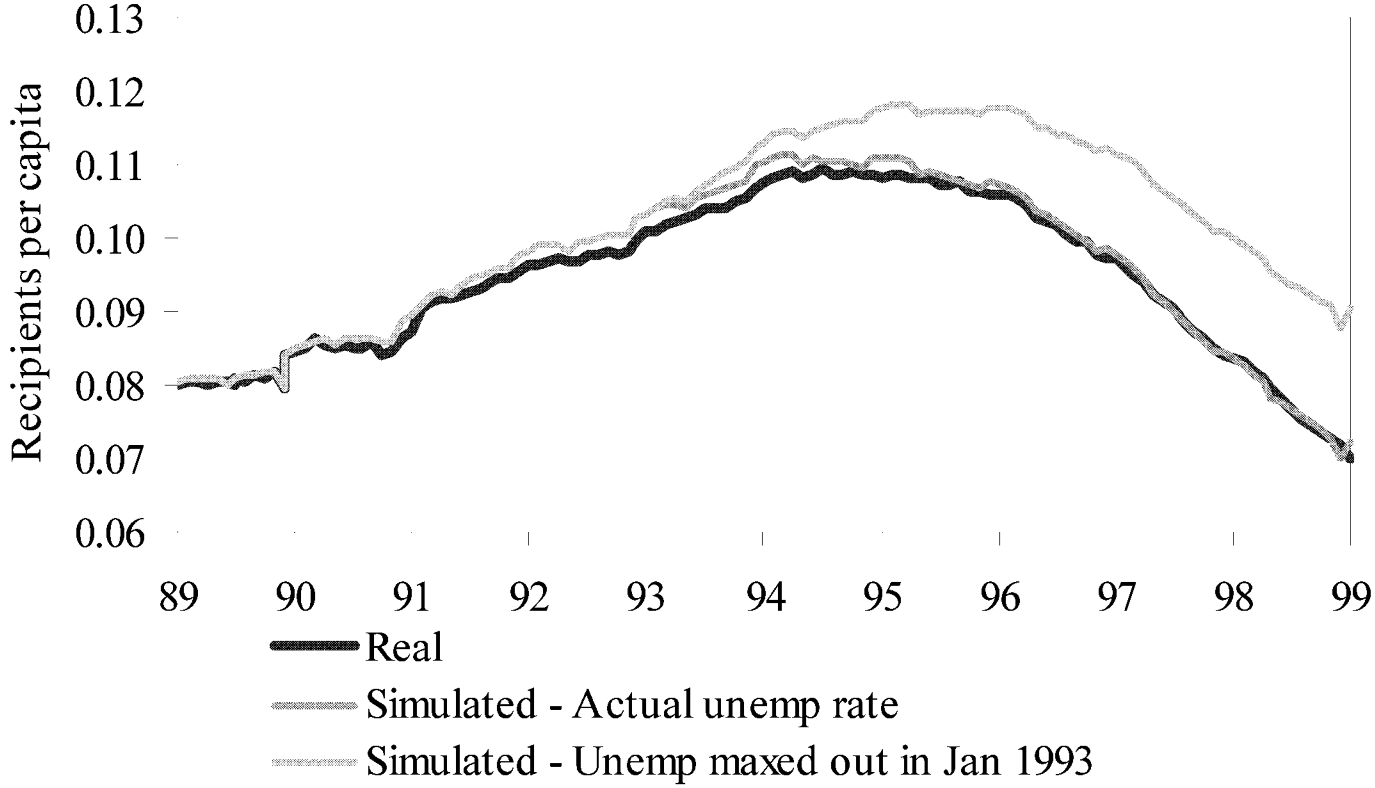
NOTE: Clean figure copies that are supposed to be “camera ready” follow. These are repeats of Figures 1, 2, and 3, respectively.



3,000,000
2,500,000
2,000,000
1,500,000
1,000,000
500,000
0
Recipients
12.0
10.0
8.0
6.0
4.0
2.0
0.0
Unemployment Rate
Jan-87
Jan-88
Jan-89
Jan-90
Jan-91
Jan-92
Jan-93
Jan-94
Jan-95
Jan-96
Jan-97
Jan-98
CA237
MEDS
Unemp



Logit Coefficient
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
3
6
9
12
15
18
21
24
27
30
33
36
39
42
45
48
Months on Aid



Recipients per capita
0.13
0.12
0.11
0.10
0.09
0.08
0.07
0.06
89
90
91
92
93
94
95
96
97
98
99
Real
Simulated - Actual unemp rate
Simulated - Unemp maxed out in Jan 1993