Migration, the lifecycle, and state benefits: How low is the bottom?

Jonah B. Gelbach*
University of Maryland at College Park

This version: October 2, 2003
First version: March 4, 2002

Abstract

I show that among women likely to use welfare, movers move to higher-benefit states. I also find that the probability likely welfare users will move at all is lower in higher-benefit states. This effect is concentrated early in the lifecycle, as theory predicts. I construct a theoretical framework to measure the impact of welfare migration on optimal state benefits. Simulation results suggest little impact in higher-benefit states, but possibly a more substantial impact in other states. Lastly, evidence suggests little reason for concern in using cross-state variation in welfare generosity to identify incentive effects of the welfare system on other outcome variables. *JEL: H73, I38, J61, R23.*

*I would like to thank the editor, an anonymous referee, Marianne Bitler, Sandy Black, Bill Evans, Jeff Grogger, Judy Hellerstein, Hilary Hoynes, Jesse Rothstein, Duncan Thomas, and participants at Berkeley, Davis, the IRP Summer Workshop, and the RAND/UCLA labor seminar for helpful comments and discussions. I am grateful for generous provision of data on cost of living differences (compiled by Steven Craig and provided by Phil Levine and David Zimmerman), minimum wages (provided by Jeff Grogger), state maximum welfare benefits (provided by Hilary Hoynes), and estimated state Medicaid expenditures (provided by Robert Moffitt).

1 Introduction

In this paper, I use data from the 1980 and 1990 Censuses to answer three questions. First, do cross-state differences in welfare generosity affect state-to-state migration by single mothers? Second, how large an impact could any such welfare migration effects have on states' optimal benefit levels? Third, if migration responds to welfare benefit levels, is it appropriate to rely for identification on cross-state variation in benefit levels in other research?

These questions arise because the modern U.S. system of cash welfare is highly decentralized. Even before recent reforms, cash benefits paid through Aid to Families with Dependent Children (AFDC) were set by the states. Recent state-level reforms, together with the federal Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA), have given states (and even some cities and counties) a great deal of flexibility in the design of their welfare systems. Advocates of such federalism praise it both because local preferences are given primacy and because many jurisdictions may be effective reform "laboratories". Critics worry that mobility inevitably induces a "race to the bottom": even otherwise-generous states may cut benefits to encourage long-term welfare participants to leave or to avoid becoming "magnets" for long-term welfare participants.

Policymakers have long worried about welfare migration. The issue was recognized explicitly in restrictions on residency requirements appearing in the Social Security Act of 1935, which created AFDC. Since then, residency requirements twice have come to the attention of the Supreme Court [*Shapiro v. Thompson* (1969) and *Saenz v. Roe* (1999)], which both times found them unconstitutional.[1] Some observers question the relevance of welfare migration, either before or after the passage of PRWORA and its lifetime time limits. But state policymakers do not appear to share this skepticism. A number of econometric studies concerning state policy determination have found significant effects of neighbors' policies on own policies (*e.g.*, Baicker (2001), Gramlich & Laren (1984), Case, Hines & Rosen (1993)). This result is frequently ascribed to concerns about migration. Moreover, state policymakers consistently *say* they are worried about welfare migration.[2]

[1] Residency requirements of more than one year were ruled out by the Social Security Act. Prior to 1935, 20 of the 45 states having public assistance programs for dependent children required applicants to have resided in the state for at least two years, 33 states required at least one year, and 9 states required at least two years of residence in the same town or county (*Shapiro v. Thompson* (1969)). For a discussion of welfare migration concerns dating to the English Labour Ordinance of 1349, see Brown & Oates (1987).

[2] For example, see arguments in *Maldonado v. Houstoun* (1997, p. 45), one of the federal cases leading up to the *Saenz* decision, or Pear's (1997) paraphrasing of the Minnesota welfare director as follows: "preliminary evidence

What is less clear is whether state policymakers *should* be worried about migration. There have been essentially three waves of social science research on this question. The early literature is reviewed in detail by Cebula (1979) and characterized by Moffitt (1992, p. 34) as "severely hampered by a high level of data aggregation" and the resulting inability to focus on groups for whom welfare migration is likely. Several authors used more disaggregated data in the 1980s, finding significant effects of welfare benefit generosity on migration in the expected directions (*e.g.*, Gramlich & Laren (1984), Blank (1988), Peterson & Rom (1989)). More recently, a consensus seems to have developed that any welfare migration effects are too small to affect rational policymakers' choices. This view is primarily the result of influential studies by Levine & Zimmerman (1999) and Meyer (1998). For example, Meyer & Rosenbaum (2001, p. 1086) cite the "small effect of welfare on migration" found in Meyer (1998) in ignoring endogeneity of location. Citing Levine & Zimmerman, Baicker (2001, p. 155) writes that "The existence of this 'welfare magnet' effect is controversial" and later refers to "lack of evidence that [welfare] migration occurs" (p. 177).[3]

One notable feature of the literature on welfare migration is the wide variety of methods authors have chosen to measure it. Some authors focus on population flows into or out of states; examples include much of the early literature as well as Walker (1994), Levine & Zimmerman (1999), and (partly) Meyer (1998). A less common approach, taken for example by Blank (1988) and Kennan & Walker (2003), is to estimate a structural, multinomial model of locational choice. A number of other approaches have been taken; for example, Meyer (1998) considers differences in welfare participation rates among movers and non-movers, while McKinnish (2003) focuses on testable implications regarding county-level welfare expenditures. The lack of consensus regarding the proper econometric approach—indeed, regarding even the proper dependent variable—is no accident. Any structural model that allows more than a trivial wage distribution and also allows more than a small number of jurisdictions is doomed to infeasibility.[4] On the other hand, using reduced form methods raises questions concerning both the proper dependent variable and how to control for characteristics of potential destination states.

In this paper I will address the dependent variable issue by using two complementary methods

indicated that migration to Minnesota had declined slightly since [a 30-day residency requirement] took effect."

[3] A recent exception to this trend is McKinnish (2003), who cleverly compares aggregate welfare expenditures in border counties with expenditures in interior counties. She finds that expenditures in counties bordering a state with monthly benefits $100 less than own benefits are 4–8% greater than for interior counties.

[4] Kennan & Walker (2003, p. 8) report that with a separate jurisdiction for each U.S. state and a 5-point wage distribution "the number of dynamic programming states is 40,414,063,873,238,203,032,156,980,022,826,814,668,800."

to measure welfare migration. First, in section 2, I consider whether women who do move, move to higher-benefit states. Other things equal, it seems intuitively reasonable that this will be true if welfare plays a role in migration decisions. The results clearly suggest that among single mothers likely to spend long periods of time on welfare, movers migrate to higher-benefit states. This result holds for both the 1980 and 1990 samples, and it is robust to the inclusion of a large number of state-level control variables. This method does have drawbacks, however. For example, it is possible in principle that women are more likely to move out of a state when it cuts its benefit level, but that they move to states with similar or even lower benefit levels when they move. This could happen if information regarding other states' benefits were poor, if state benefits were geographically correlated, or if higher-benefit states tended to have lower levels of wages or other amenities valued by welfare mothers.

A more serious problem with studying only conditional move-up behavior is that it may have little bearing on the key policy question raised in the introductory paragraph: how large an impact could welfare migration effects have on states' optimal benefit levels? Policymakers who seek to push welfare mothers out of their states should not care whether those women move to higher- or lower-benefit states: what matters is simply that they move out of the state at all.[5] I address this point in section 3 by estimating probit models in which the dependent variable indicates whether a woman's state of residence in the Census year differs from her state five years earlier, while the right-hand side variables of interest involve state welfare benefit generosity. If welfare migration is an important enough phenomenon to justify a race to the bottom, then policymakers must think that women will be more likely to move out of a state with lower benefit levels. This question is most directly addressed by relating out-migration behavior—wherever the destination—to the benefit level in women's state of residence five years before the Census.

A key feature of these probit models is that I allow the effect of welfare benefit generosity on migration to vary with children's ages. Migration to a higher-benefit state is more advantageous to women with young children, because they can receive benefits for a longer period of time. If potential welfare participants are forward-looking, then they will place at least some value on the stream of benefits available in the future. I call this effect the dynamic incentive effect.[6] Dynamic

[5] Similarly, policymakers should not care whether women moving into their states came from higher- or lower-benefit states, though I do not address this question empirically here. In section 4, I argue that focusing on out-migration is unlikely to greatly alter my qualitative conclusions.

[6] This effect has not generally been noted in the welfare migration literature; an exception is an earlier paper of

rationality is not an uncontroversial assumption for the welfare population,[7] but because of a dynamic selection effect, lifecycle migration effects should exist even in the absence of dynamic rationality. Imagine that a woman with only a one-year horizon would find it beneficial to move this year. Then other things equal, she will not move next year, since she has already done so this year. Such selection effects imply that welfare migration will be concentrated earlier in the lifecycle, even among totally myopic women. Because welfare eligibility does not start until a woman has a child, measures of children's age appropriately index the lifecycle.[8]

The probit results on out-migration suggest that for the 1980 Census sample, welfare benefit levels play a substantial role in state-to-state migration decisions among single mothers most likely to spend extended periods of time on welfare. Among women likely to spend less time on welfare, there is considerably less (and sometimes zero) evidence of welfare migration. The probit results for the 1990 sample are much less clear. Significant coefficient estimates occur not only for never-married dropouts, but also for other groups of women whose long run welfare use is likely to be very limited. In some cases, the estimates for these comparison groups suggest greater welfare migration than for the never-married dropouts. The most reasonable conclusion for the 1990 probit estimates are thus that they show no uncertain evidence of welfare migration.

In section 4 I construct a simple model of optimal state policy determination in which each state policymaker takes other states' benefit choices as given and then chooses its own benefit to maximize a social welfare function. This model suggests the form of a theory-driven, simulation-based metric for the importance of welfare migration to policymakers, which I then implement in section 4. Introducing such a metric is important since different authors measure welfare migration in different ways, and the implications for policymakers frequently are not clear. By contrast to much of the empirical literature on welfare migration, my methodology focuses on marginal costs, rather than (either marginal or average) person counts. The end result is a rough estimate of the ratio of optimal state-level welfare benefit levels in a world with no welfare migration, to optimal benefits given observed levels of welfare migration. Because the 1990 probit results are mixed, I

mine, Gelbach (2000), in which I develop the issue of dynamic incentive effects in detail.

[7] To be sure, recent work on time limits by Grogger & Michalopoulos (2003) and Grogger (2000) suggests that welfare participants are at least partly foward-looking. Another prominent example is Hubbard, Skinner & Zeldes (1995), who examine effects of social insurance programs on savings behavior of the poor using a dynamic programming model with four state variables. They simply assume full dynamic rationality, as do Kennan & Walker (2003). Ziliak (2003) also considers wealth accumulation in a dynamically rational framework.

[8] As I discuss in Gelbach (2000), the dynamic incentive and selection effects are isomorphic to the phenomena of state dependence and baseline heterogeneity, discussed widely in the literature on survival models.

carry out the simulation only for the 1980 sample.

The simulation results suggest little basis for concern that welfare migration reduces optimal state-level welfare benefit levels in higher-benefit states. For example, under a log utility assumption, my upper-bound point estimate for Wisconsin suggests that optimal benefits are reduced by only 5.4 percent as a result of welfare migration. These estimates do not account for in-migration effects, which might be particularly large in higher-benefit states, but I argue below that even in extreme circumstances, accounting for in-migration would increase my estimates by less than a factor of two. For lower-benefit states, estimated effects on optimal benefit levels are also small provided I assume that welfare migrants are no different from other women in their use of public assistance income. But the results are quite different if I assume that welfare migrants would receive the maximum available welfare benefit—which seems reasonable given that such women change migration behavior because of welfare generosity. The point estimates under this upper-bound assumption suggest that in the absence of welfare migration, optimal benefits would be 36.5 percent greater in Mississippi (a very low-benefit state) and 9.5 percent greater in Illinois (approximately the median-benefit state). These estimates suggest some cause for concern about the potential effects of welfare migration on optimal state benefit levels for lower-benefit states.

The third question raised above concerns whether welfare migration invalidates researchers' use of state benefit levels as independent variables in studies of welfare's other potential incentive effects. In section 5 I show that, and explain why, the clear answer is no. I summarize the findings and discuss the importance of recent welfare reforms in section 6.

2 Do Women Likely to Use Welfare Move Up?

As a simple first pass at determining whether women who move, move to higher-benefit states, consider the first column of Table 1. Panel A reports that 66% percent of 1980 never-married dropouts had some public assistance income in 1979.[9] Panel B reports the fraction after subtracting the rate for women of other marital history/educational attainment. Thus the welfare participation rate for never-married dropouts relative to ever-married highschool graduates (whose welfare participation rate in 1979 was 17%) is $66\% - 17\% = 49\%$. Results for groups of women in Panel B are reported in order of welfare participation rate, so that as we move down the column we consider comparison

[9] I discuss a number of issues related to data construction in Appendix A.

groups with lower welfare participation rates.

This column shows that never-married dropouts are much more likely to use welfare in any given year than are the other groups of women. In fact, it is interesting to note that welfare use among never-married dropouts is as different from ever-married dropouts as it is from never-married highschool graduates. Bane & Ellwood's (1994) simulation evidence reinforces this result when long run welfare use is considered. They simulate dynamic models estimating both the expected time (within a 25-year window) a woman will spend on AFDC and the probability she will spend at least 10 years on welfare. Because welfare migration makes much more economic sense when a woman anticipates long-term welfare receipt, such dynamic considerations are a key element in choosing data groupings. Bane & Ellwood find that at the time she starts a first welfare spell, a never-married dropout can expect to spend more than 10 years on AFDC, and she has a 45% chance of spending at least that many years on welfare. By contrast, an ever-married highschool graduate who begins a first spell on AFDC can expect to spend just over 4 years on welfare, with her probability of spending at least 10 years on AFDC being only 10%. Bane & Ellwood's figures for never-married graduates and ever-married dropouts are similar to each other, suggesting that if they start a welfare spell, these women can be expected to spend about six years on welfare. These statistics are conditional on starting a spell, and Table 1 clearly suggests that never-married dropouts are much more likely to start a spell than are all other groups of women. For this reason, I focus on never-married dropouts as the main "treatment" group throughout the paper.

In the second column of Table 1, I report the fraction of never-married dropouts who moved to a new state between 1975 and 1980. Just under five percent of never-married dropouts moved between 1975–80. Panel *B* shows that this rate is significantly lower than migration rates for other women, with a difference ranging between two and six percentage points. This finding is not surprising given that moving is costly, and that never-married dropouts generally have lower lifetime resources than do other women. Of course, the question of interest is not whether never-married dropouts are more or less likely to move, but rather the relationship between migration and state welfare generosity.

To see whether never-married dropouts who do move are moving to higher-benefit states, in the third column of Panel *A* I report the change in the state maximum monthly benefits that is due to migration among never-married dropouts. To construct this change, I subtract the state maximum welfare benefit in a woman's 1975 state of residence from the maximum benefit in her

1980 state. For never-married dropouts who moved, average monthly benefits were \$27 greater in destination states than in origin states, so these women did move to higher-benefit states on average.[10] Panel B subtracts the change in maximum benefits for other groups of migrating women from the change for never-married dropouts. For example, the average monthly maximum benefit level available to migrating ever-married graduates fell by \$16. Thus the difference-in-differences estimate for never-married dropouts compared to ever-married graduates is $\$27 - (-\$16) = \$43$. The difference-in-differences estimates are correctly signed for all five comparison groups, ranging from \$25 to \$54.

Simply looking at the change in average benefits among movers might be misleading, however. As an extreme example, suppose never-married dropouts were all initially located in the lowest-benefit state. Then those who moved would necessarily move to higher-benefit states, even if welfare benefits had nothing to do with their motivation for moving. To solve this problem, in the fourth column I present the change in benefits for never-married dropouts as well as difference-in-differences estimates conditional on the maximum benefit in the woman's 1975 state of residence. The figure reported in Panel A is the estimated constant in a regression that includes only never-married dropouts, with the lone non-constant covariate being the maximum benefit in the 1975 state of residence. Because the maximum benefit variable is de-meaned, the constant estimates the expected change in benefits for never-married dropouts under the counterfactual that they all started out in a state with the mean maximum benefit level. The conditional change in maximum benefits for never-married dropouts when the state-of-origin maximum benefit level is included is \$25. This suggests that never-married dropouts who move tend to live in disproportionately lower-benefit states, but clearly the basic conclusion is unaffected. Looking at the relative changes for other groups of women in Panel B when initial maximum benefit level is held constant again shows that never-married dropouts who move are systematically moving to higher-benefit states.

The table's final column again includes the 1975 maximum benefit level, while seeking to control for other baseline differences by adding a set of demographic and two sets of state-level controls. The demographic controls are: a dummy for being white; a dummy for being Hispanic; a quadratic in mother's own age; a cubic in youngest child's age as of five years before the Census; a cubic in oldest child's age as of five years before the Census; the number of children ever born as of five years

[10]Mean values of the maximum available benefit are reported in an appendix table, Table 8. For never-married dropouts, the mean in their 1975 state of residence was \$561 per month.

before the Census; and the interaction of all youngest child's age terms with number of children ever born as of five years before the Census.

The first set of state-level controls consists of variables concerning the woman's 1975 state of residence. The own-state variables are: *per capita* income level, the larger of the state and federal minimum wages, and the state's unemployment rate (these variables would enter a reduced form wage equation); estimated Medicaid expenditure for an AFDC family of four; per-student schooling expenditures; and a state-level measure of the ratio of "marriageable men" to women (this variable is essentially the ratio of high-income men to women within given demographic groups).

The second set of state-level controls concerns characteristics of states located near the woman's 1975 state of residence. How to deal with the alternative states of residence besides the state of origin is an important question in a migration study. One reasonable approach might be to control for characteristics of states that share a border with state j.[11] The problem with this approach is that the number of border states varies systematically across regions of the United States. Because the most clustered states are in the Northeast, and because these states have relatively generous benefits, border states may systematically under-represent nearby alternatives in higher-benefit states.

As an alternative approach, I construct a 500-mile region for each state. I use Census data on latitude and longitude for the population centroid of each U.S. place to calculate the minimum border-to-border distance for each pair of states. For each state, I then construct a list of other states whose borders are within 500 miles; these states constitute the 500-mile region for the initial state.[12] All variables used as own-state controls (discussed above) have at least one regional counterpart. To understand how I construct these variables, consider first the regional maximum benefit. For each state j, I find the maximum benefit available in any other state having a border within 500 miles of j, which then becomes the value of state j's regional welfare benefit level. Notationally, let the set of states in state j's region be $r(j) \equiv \{k : \text{State } k \text{ is within 500 miles of } j\}$. Then state j's regional benefit value is $b_{r(j)} \equiv \sup_{k \in r(j)}\{b_k\}$. I use the same supremum approach

[11] For example, see Kubik & Meyer (1992) (cited in Meyer (1998)) and Walker (1994).

[12] A table listing the region for each state is available on request. The 500-mile regions often include many other states. Washington state's seven members (California, Idaho, Montana, Nevada, Oregon, Utah, and Wyoming) is the least, while Indiana, Kentucky, Michigan, and West Virginia all have 29 regional members. Preliminary work using two sets of regional variables—both the ones used here and variables based on 200-mile regions—were highly imprecise. There is already substantial collinearity between own-state variables and the 500-mile regional controls, so I focus only on the 500-mile regions.

to generate a regional value for all other state-level control variables except the unemployment rate. Because higher unemployment rates reflect both the number of people searching (including recent labor force entrants) and the difficulty any particular labor force participant has in finding a job, I include both the maximum and minimum values of the unemployment rate within a state's region.[13]

The final column of Table 1 adds the demographic and state-level controls. Again the key fact is that the estimates do not change much with the addition of controls. In Panel B, we see that controlling for the demographic and state covariates does not change the main conclusion drawn from the third column; if anything, adding the controls strengthens the results. Table 2 replicates Table 1 using data from the 1990 Census. The particular numbers differ, and the extent of apparent welfare migration seems to have dropped, but the 1990 results are qualitatively similar to those for 1980.

The clear conclusion from Tables 1 and 2 is that never-married dropouts who move, move *up*. This is true whether we consider them in isolation or in comparison to several other groups of women for whom we should expect less welfare migration. These results provide firm evidence suggesting that the direction of migration among never-married dropouts is consistent with the welfare migration hypothesis.[14] However, as noted in the introduction, this method is not directly tied to the question of whether race-to-the-bottom incentives are empirically important for policymakers. For this reason, I now turn to the relationship between welfare benefit levels and whether women move out of state at all.

[13]While the approach taken here is to my knowledge more comprehensive in its use of regional controls than any other in the welfare migration literature, a variety of other approaches have been taken in other empirical literatures. For example, in their study of hospital competition, Kessler & McClellan (2000) measure the attractiveness of hospital j to person i along dimension h (for example, number of beds) by taking the distance from i's residence to hospital j, minus the distance from i's residence to the nearest hospital j' having the same value of characteristic h as hospital j; they also compute the distance to the nearest j'' with a different value of characteristic h. This approach is similar in spirit to mine. Case et al. (1993) use a more formal spatial model in estimating the impact on a given state's spending policies of other states' policies. Their approach involves deriving similarity weights (along dimensions including geography, income level, and racial composition), which are used create indices of other states' spending patterns. They then regress own-state spending on this index. This index approach would be fruitful in the present context if we knew how to choose weights for each individual and each state. But in the context of migration decisions, this choice of weights is a function of the output of a fully structural model, which is computationally infeasible without some set of ad hoc assumptions that reduce the state space (e.g., see Kennan & Walker (2003)).

[14]As an additional check, I created tables analogous to Tables 1 and 2 in which the focus is on a binary variable indicating whether a migrating woman moved to a higher-benefit state, rather than the change in available benefits. Results in these tables are qualitatively similar to those reported here and are available on request.

3 Probit results for out-migration

This section concerns probit results in which the dependent variable is a dummy indicating whether a woman moves out of state at all. It also addresses the dynamic incentive and selection effects discussed in the introduction by interacting the maximum benefit variable with oldest child's age as of the beginning-of-period year (either 1975 or 1985). Oldest child's age can be observed for a subset of women in the Census, because the data include ages for all persons in the household, the relationship to the household reference person, and the number of children ever born for every woman in the Census. I thus restrict attention to only those women for whom all children live in the household.[15] In the presence of welfare migration, the sign on the "main" effect for the maximum benefit variable should be negative: raising benefits reduces the probability of out-migration for a woman whose oldest child is newborn. Lifecycle welfare migration effects, whether due to incentives or selection, are present in the expected direction if the coefficient on the interaction term is positive: the benefit effect decreases in magnitude as a woman's oldest child ages.

Migration presumably involves long run economic concerns, and state of residence is not randomly re-assigned each period. Thus variables that index the lifecycle for purposes of welfare migration are likely to be associated with migration behavior, even holding welfare benefit levels constant. To avoid wrongly attributing this sort of lifecycle variation in migration to lifecycle welfare migration via the oldest child's age interaction, I include a number of age-related control variables in the probits. These variables are a quadratic in mother's age and cubics in both oldest and youngest child's age as of five years before the Census. Birth spacing and total fertility may be endogenous to welfare use, so I also include a linear term in the number of children born as of five years before the Census and interact this variable with the cubic in youngest child's age. Also, I interact all state-level variables discussed in section 2 with oldest child's age.

3.1 Probit estimates allowing for lifecycle variation

I report estimates for the 1980 sample in Table 3. The first row contains the coefficient on the maximum benefit main effect, while the second row reports the coefficient on the interaction of

[15] I use oldest, as opposed to youngest, child's age for several reasons. While youngest child's age is more closely tied to the dynamic incentive effect, oldest child's age is more closely related to the dynamic selection effect, and it is also highly correlated with youngest child's age. Also, I do not observe completed fertility for most women, so observed youngest child's age is a noisy measure of planned last child's age. Lastly, fertility may be endogenous to welfare benefit levels, whereas oldest child's age may be treated as predetermined once a woman has any children.

the maximum benefit level and oldest child's age as of 1975. Thus the main effect represents the impact on the probit index of changing the monthly welfare benefit by \$100 for a woman whose oldest child was age-0 in 1975. The interaction coefficient represents the change in that impact when we consider a woman whose oldest child is one year older (other things equal). Consider first the estimates for 1980 never-married dropouts in the first column of the table. The benefit variable is scaled so that the reported coefficient represents the estimated impact on the probit index of increasing the monthly benefit by \$100. Thus the estimated column [1] age-0 coefficient estimate of -0.094 for never-married dropouts implies that a \$100 increase in the maximum benefit level reduces the probit index by (an insignificant) 0.094 standard deviations of the probit residual.

In column [2], I add the own-state covariates mentioned in section 2, while column [3] also includes the regional controls. The coefficient estimates provide strong evidence of lifecycle welfare migration for never-married dropouts. The age-0 coefficient is negative and significant in both columns that include these variables, while the interaction coefficient is positive and significant in these columns as well. I note that the significant role of welfare benefits would have been missed without including at least the own-state controls.

Probit coefficients cannot be compared directly across specifications, because including the state-level variables in the model reduces the standard deviation of the probit residual term. Even if the added controls and the state maximum benefit variable are independent, the maximum benefit coefficients will increase in magnitude as long as the added controls have some explanatory power, sincd the coefficient estimates have units equal to the standard deviation of the unobservable. Fortunately, marginal effects are not subject to this scale indeterminacy, as they are computed after integrating over all variables that are either truly unobservable or observable but omitted from the estimation. Thus marginal effects may be meaningfully compared across specifications.[16] The next two rows of Table 3 report marginal effects for women whose oldest child was aged 0–5 and 6–11, in 1975. For any mother i, I compute the estimated marginal effect as $(\hat{\beta}_0 + a_i\hat{\beta}_\times)\phi(\widehat{z}_i)$, where $\hat{\beta}_0$ is the estimated maximum benefit main effect, a_i is the age of mother i's oldest child, $\hat{\beta}_\times$ is the estimated coefficient on the benefits-oldest child's age interaction term, $\widehat{z}_i$ is the estimated probit index for mother i, and ϕ is the standard normal density function. I then average over the sample of women whose oldest child is aged 0–5 or 6–11 to get the reported marginal effects. The

[16] For a more formal and detailed treatment of interpreting probit coefficients across specifications, see the excellent discussion in Wooldridge (2002).

presence of the probit index means that the variance in the marginal effect will depend on the variance of all coefficient estimates; reported standard errors are computed using the delta method and take account of this fact.

The column [3] estimates for never-married dropouts whose youngest child was 0–5 in 1975 show that on average, a marginal change in the maximum benefit level would reduce outmigration by 2.4 percentage points. This is a very large effect—43%—given that the baseline outmigration rate for these women (shown in the next row) is only 5.6 percent. By contrast, the estimated marginal effect for women whose oldest child was aged 6–11 in 1975 is virtually 0. Thus there is very strong evidence of lifecycle welfare migration for never-married dropouts in the 1980 sample.

The results for ever-married single dropouts in the 1980 sample show essentially no evidence of welfare migration. The ever-married graduates' implied marginal effects are significantly different from zero for oldest children aged 6–11 (though not generally for oldest children 0–5). However, both estimated marginal effects are small relative to the baseline outmigration rates. For 1980 never-married graduates, the probit coefficient estimates are significant only when the state controls are omitted. However, the implied marginal effects are significant when both the own-state and regional controls are included. Significance of the marginal effects when the coefficient estimates are insignificant is unsurprising for these women: the two probit coefficients of interest covary, and they are jointly significant ($p \leq .04$) in the column [3] specification. In any case, the estimated third-column marginal effects for never-married graduates are significant and negative regardless of oldest child's age; the marginal effect for the women with older oldest children is slightly larger, though the difference is not statistically significant. Compared to the baseline outmigration rate, these estimates are nontrivial, though they are smaller in both absolute and relative terms than the estimated effect for never-married dropouts with younger oldest children.

I report analogous estimates for the 1990 sample in Table 4. The results again suggest statistically significant effects of welfare benefits for never-married dropouts, and these effects are consistent with the lifecycle hypothesis. However, almost the same could be said about both groups of highschool graduates, regardless of marital history. The absolute size of the mean marginal effects for highschool graduates with oldest children aged 0–5 is considerably larger than that for the never-married dropouts with similarly aged children; the relative effect size is approximately equal across the three groups. These results suggest caution for the 1990 sample. It might be wrong to conclude that there is no evidence of welfare migration, but the presence of significant coefficients and marginal effects for ever-married graduates suggests that something other than welfare migration is at play here as well.

Overall, these results suggest three basic conclusions. First, for the 1980 sample, there is strong evidence of lifecycle welfare migration among never-married dropouts, the group most likely to spend extended periods of time on welfare. Second, estimated welfare migration effects among other 1980 groups are smaller, in some cases much smaller. Third, the results for the 1990 samples are too mixed to allow a definitive conclusion in either direction.

3.2 Robustness checks

3.2.1 Selection and marital history

Once a woman marries she can never again be never-married, so the probability of being never-married necessarily falls as the lifecycle progresses. Since my focus is partly on lifecycle aspects of welfare migration, it is important to make sure my results are not spuriously driven by endogenous selection into ever-marriage. The women who would cause this sort of selection are those who are likely to marry and then later divorce, so that they enter my sample of ever-married single mothers. One approach to addressing this concern is to estimate models that pool over marital history while still stratifying on educational attainment. Results using that approach, reported in Gelbach (2002), show significant evidence of lifecycle welfare migration for 1980, as here.

A second, more formal, approach is to estimate the models reported above jointly with a selection equation. This allows for estimates that account for correlation in the unobservable factors affecting the probability of outmigration on the one hand, and being in a given marital history–educational attainment cell on the other. For purposes of testing for selection, I take the most common parametric approach and assume bivariate normality of the two unobservables. To construct instruments for the selection equation, I first take all women in the PUMS whose first birth occurred in the Census year or the preceding year. I then calculate the fraction of these women in each state who fall in each of three marital history/status cells (denoted mh) and three educational attainment cells (denoted e). The cells are defined by (i) being married at Census time, single and never-married, or single and ever-married; and (ii) being a highschool dropout, a high school graduate without a college degree, and being a college graduate. These instruments thus capture the propensity $z_{mh,e}$ of recent new mothers to fall in each cell, by state.

One would expect the state-specific composition of marital history/status and educational attainment of new mothers to reflect unobservable state-level factors (*e.g.*, norms) that also affect women with older children, and I do find that the instruments are jointly highly significant in the selection equations. It is reasonable to think these instruments also satisfy the exclusion restriction for instrument validity. None of the women used to construct the instruments is included in my outmigration sample, since that sample includes only those whose first child had been born five years before the Census. Also, the large number of included state-level controls suggests that all of the obvious confounders are already in the models. Lastly, to my knowledge, no one has so far argued that variables like my instruments belong on the right-hand side of a migration equation; certainly, no one has *included* such variables. Thus the literature as it stands would seem to sanction the use of these instruments.

I thus re-estimated the eight third-column models in Tables 3 and 4 using these instruments in a bivariate probit model with selection. The selection term was significantly different from zero for all 1980 specifications except for the never-married dropouts. However, the estimates for the coefficients of interest—those relating to the state maximum benefit—are essentially unchanged for all four specifications. Most importantly, the never-married dropouts results continue to be correctly signed and statistically significant, while none of the other samples yield significant estimates. For the 1990 results, the selection correction term is never significantly different from 0, and the basic relationship between maximum benefit levels and migration is similar to that in Table 4.

3.2.2 Cross-state price differences

As noted in Appendix A, the state-level variables used in the main results have been adjusted only for national price changes over time. While one would expect long run differences in state-specific economic variables to capitalize different values of local amenities, welfare mothers are unlikely to be the marginal consumers for whom utility is equalized across state of residence. Thus one concern is that my results do not take adequate account of cross-state differences in cost of living. In Gelbach (2002), I used a cross-state index to adjust for price differences.[17] When I re-estimate

[17] Cost of living data for the period 1973–1993 were compiled by Steven Craig and were also used in Levine & Zimmerman (1999). The index numbers were constructed using region-wide rural price indices (through 1990), as well as city-specific budget studies conducted by the Bureau of Labor statistics for several years in the 1960s and 1970s. These figures were then updated for later years using city-specific inflation rates provided by the Census Bureau. State-level numbers are weighted averages of the urban and rural numbers, with weights equal to within-state population share.

the models reported here while using that index, none of my qualitative conclusions change.

3.2.3 Single college graduates and married mothers

Single college graduates and married mothers are another possible comparison group, since they have relatively low long run welfare use; both Meyer (1998) and Levine & Zimmerman (1999) use married mothers in this fashion. I estimated models analogous to those reported above for these women, with results reported in Table 5. All right-hand side covariates used in the previously discussed models are used here as well. For the 1980 sample, these robustness checks provide little basis for concern: virtually all the estimated probit coefficients for married mothers in the 1980 sample are statistically insignificant and small. For 1980 single college graduates, the age-0 effect is statistically significant in the second two specifications, but these coefficients are wrong-signed.[18] For the 1990 samples, the results are again mixed. For example, the age-0 effect is significant and negative for all three specifications using married highschool graduates; it is significant and negative for two of the three specifications using married college graduates. This pattern provides further reason to think the maximum benefit variable may be picking up some spurious factor for the 1990 sample. Lastly, I note that estimating these specifications using the cost of living index to adjust state-level variables does not change these substantive conclusions.

4 Estimating the welfare migration impact via simulation

In subsection 4.1, I briefly describe a theoretical framework used to study the effect of welfare migration on optimal benefit levels. While the basic ideas are straightforward, a precise exposition is lengthy and thus confined to Appendix B. In subsection 4.2, I provide a heuristic discussion of the procedure used to simulate the key elements of the theoretical framework; the details appear in Appendix C. I then discuss estimates of the impact of welfare migration on optimal benefit levels. Finally, in subsection 4.3 I map results from Levine & Zimmerman (1999) and Meyer (1998) into my framework, showing that their estimates imply considerably smaller effects of welfare migration on optimal benefit levels.

[18] I do not stratify single college graduates by marital history because there are very few never-married single mothers who have at least 16 years of education (for 1980) or at least a Bachelor's degree (for 1990). A referee points out that education may be ongoing for a substantial fraction of women younger than, say, age 25. I thus re-estimated the models reported for single college graduates including only single mothers aged 25 or older. The results echo those in Table 5.

4.1 A theoretical framework

Assume that state policymakers seek to maximize a social welfare function subject to a budget constraint. This is a standard approach taken in both empirical and theoretical papers studying normative policy design (*e.g.* Akerlof (1978), Mirrlees (1971), Moffitt, Ribar & Wilhelm (1998), Baicker (2001)). To fix ideas, suppose that a given state policymaker's problem is to

$$\max_{\{b,z,t\}} u(b) + F(z,t) - \lambda G(z,t,b), \tag{1}$$

where

$$u(b) \equiv \begin{cases} \frac{b^{1-\rho}}{1-\rho}, & \text{if } \rho \neq 1 \\ \ln b, & \text{otherwise} \end{cases} \tag{2}$$

The first part of the social welfare function implies that welfare benefits enter the policymaker's preferences via constant relative risk aversion (CRRA) utility over the income received by a representative woman with no income other than from welfare. For concreteness, I will assume that $\rho = 1$, i.e. log utilities. The function $F(z,t)$ is meant to capture the utility of all non-welfare recipients given non-welfare government expenditures z and taxes t (the relative weight placed on non-welfare recipients is implicit in this function). The Lagrange multiplier on the budget constraint is given by λ, and the budget constraint is represented by the net expenditures function $G(z,t,b)$. If labor supply effects of welfare benefits are not too great for the state economy as a whole, we may write $G(z,t,b) = C(b) + Q(z,t)$, so that welfare-related expenditures are additively separable from revenues and other expenditures. We can express $C(b)$ as $S(b)N(b)$, where $S(b)$ represents the state's welfare expenditures per potential welfare recipient and $N(b)$ represents the state's population of potential welfare recipients.

The key counterfactual in deploying this framework is what I call the "no welfare migration world", in which state welfare expenditures are somehow independent of migration behavior. Define the optimal welfare benefit levels in the observed and no welfare-migration worlds as b^* and b^{**}, respectively. The discussion in Appendix B provides plausible conditions under which

$$R \equiv \frac{b^{**}}{b^*} \sim 1 + \frac{S(b^*)N'(b^*)}{S'(b^*)N(b^*)} = 1 + \widetilde{r}, \tag{3}$$

The ratio $\widetilde{r}$ is the object of interest here, because it represents the ratio of optimal benefits in the no welfare migration and observed worlds, minus 1. When $\widetilde{r}$ is substantially greater than 0, there is cause for concern about a race to the bottom. The numerator of $\widetilde{r}$ is what I call the "welfare migration effect". I call the denominator of $\widetilde{r}$ the "constant-population effect." Holding constant the state's population, an increase in generosity increases spending because payments to incumbent welfare participants increase. This framework emphasizes two facts in assessing the policy relevance. First, it is not simply the change in population that matters, but rather the *cost* associated with this change; this is captured by the welfare migration effect. Second, it is the size of the welfare migration effect relative to the constant-population effect that determines the impact of welfare migration on optimal benefits. This is because in the observed world, both effects occur, but in the no welfare migration world, only the constant-population effect enters the first-order condition.

The main task in implementing this framework empirically is to simulate the welfare migration and constant-population effects. The probit models estimated in 3 do not allow me to relate welfare benefit generosity to migration into a state, so it is important to see whether my focus on out-migration is likely to affect my conclusions. I show in Appendix B that $\widetilde{r}$ may be written as $(1-\theta)\widetilde{r}_I + \theta\widetilde{r}_H$. Here I denotes the in-migrant population, while H denotes the "holdover" population: those women who lived in the state last period and did not move away. The weight θ is the fraction of the total constant-population effect that is accounted for by holdovers, with the balance $(1-\theta)$ being the fraction accounted for by in-migrants. Because in-migrants are such a small part of the population, θ will tend to be quite large, most likely at least 0.9. It follows that $\widetilde{r}_I$ can be much larger than $\widetilde{r}_H$ without greatly affecting $\widetilde{r}$, provided that $\widetilde{r}_H$ is not too large. For this reason, my focus on estimating $\widetilde{r}_H$ may be informative even without any results for $\widetilde{r}_I$.

A second complication of the basic framework is to incorporate the lifecycle. Appendix B one can simply estimate the numerator and denominator of $\widetilde{r}_H$ as the sum over the lifecycle of age-specific values of the welfare migration and constant-population effects. This is what I do in the simulation discussed in the next subsection.

4.2 The simulation procedure and results

In this subsection, I briefly describe my simulation procedure for the welfare migration and constant-population effects; I provide mathematical details in Appendix C. Because the probit results for outmigration are inconclusive for the 1990 sample, I carry out the simulation for the 1980 sample only. I start by taking all women who gave birth in 1980, whether married or single. I then estimate expected expenditures on a woman with given demographic characteristics when her oldest child is age a, for each $a \in \{0, 1, 2, \ldots, 11\}$. I allow marital history and status (but not educational attainment) to vary over this 12-year lifecycle in order to account for the fact that women are more likely to have been married when their children age. For each oldest child's age a, I estimate the probability that a woman is never-married as the sample mean (using the set of all women whose oldest child is aged a as of Census time) who are never-married at Census time. The probabilities of being currently married and of being ever-married and single are estimated similarly.[19]

I use the probit coefficients from the column [3] estimates in Table 3 to predict the probability of moving out over the five-year period starting when the oldest child's age is a, given other characteristics. I then calculate the implied average one-year outmigration rate, $\widehat{m}_a$. The probability of remaining in-state at oldest child's age a is $\widehat{H}_a = \widehat{H}_{a-1}(1 - \widehat{m}_a)$, with $\widehat{H}_{-1} \equiv 1$.[20] Similarly, the impact of the cut on the holdover population, ΔH_a, is estimated using the estimated difference in $\widehat{H}_a$ caused by a 10% benefit cut. In predicting five-year migration rates, I adjust mother's age and all interaction terms to take appropriate account of the given point in the lifecycle. For example, I set own age to 30 for a woman who gave birth at age 20 when the simulated value of her oldest child's age is 10. To account for additional fertility, I also add new "births," i.e., increment the number of children ever born, when oldest child's simulated age reaches each of three and five. All variables related to fertility and youngest child's age are then updated appropriately.

To estimate expected welfare spending given a woman's simulated characteristics, I use Powell's (1984) censored least absolute deviations (CLAD) estimator, since public assistance income cannot be negative. I use CLAD rather than Tobit estimation because the CLAD estimator is robust to many forms of non-normality that could cause the Tobit estimator to be seriously biased. The

[19] In results not reported here, I carried out the simulation using probit models to predict these probabilities. Because the results were qualitatively similar, I focus on the simpler approach described here.

[20] The recursive nature of the equation for $\widehat{H}_a$ implies that a benefit change that causes a woman to move out of state will reduce welfare spending for every subsequent year. The importance of this point is evident in Figure 1, discussed in section 4.3.

CLAD models are estimated separately for samples of never-married dropouts, never-married graduates, and ever-married dropouts. Given their relatively low use of welfare, I set welfare income to 0 for ever-married high school graduates, all college graduates, and all married mothers. The CLAD models include the following right-hand side variables: dummies for being white and for being Hispanic, mother's age at first birth, mother's age squared, a cubic in the ages (as of the year before the Census), of both the oldest and youngest child, children ever born as of the year before the Census, and the interaction of the youngest child cubic with children ever born. Census year values of the following state-level controls are also included: state per capita income, the state unemployment rate and the state minimum wage, as well as interactions of oldest child's age with each of these variables. I estimate the CLAD models using single mothers with oldest children aged 0–11 in the year before the Census, since that is the year to which the public assistance income variable refers. The role these estimates play is explained in detail in the appendix and briefly in the main text below.

For each woman in the simulation sample, predicted values from each of the probit and CLAD models are constructed for every oldest child's age a between 0 and 11. These values are used to construct expected expenditures given the actual state characteristics in each of three different U.S. states: Mississippi, Illinois, and Wisconsin, which are chosen to represent the bottom, middle, and top of the maximum benefit distribution. The maximum monthly benefit levels for a family of size two in these states for my 1980 data are, respectively: \$310, \$559, and \$728. Having calculated baseline expected expenditures given actual values of the state maximum benefit level, I then calculate expected expenditures when each state's maximum benefit level is reduced by 10%.

Given the estimated total spending figures, I calculate two separate estimates of the welfare migration and constant-population effects in (9) in appendix Appendix C. The first approach is to predict the CLAD index and then left-censor it at 0, yielding predicted welfare income values $\widehat{y}_i$ that can be used in equations (10) and (11) in the appendix. This approach involves assuming that a woman who is caused to alter her migration behavior when the benefit is cut would have the same expected welfare income as a woman with identical characteristics who is not caused to move. One would expect welfare migration to be more common among women with unobservably high expected welfare income, so this approach provides a lower-bound estimate of the welfare migration effect.[21] An upper-bound estimate may be obtained by replacing the predicted welfare income that

[21]The simulation estimates all account for expected Medicaid spending, based on the variable described in foot-

multiplies $\Delta\widehat{H}$ in the calculation of the welfare migration effect with the maximum benefit available (given the woman's simulated number of children). If welfare migrants are long-term welfare users, as we would expect, the upper-bound approach will provide more useful information regarding optimal welfare benefits.

Given the censoring involved in predicting the welfare migration and constant-population effects, there is little reason to believe that the sampling distribution of the $\widetilde{r}_H$ estimator is asymptotically normal; nor is the delta method appropriate even if it is. I thus use 1000 nonparametric bootstrap replications of the entire simulation procedure to estimate the quantiles of the sampling distribution of $\widetilde{r}_H$. Certain convergence issues arise in implementing this bootstrap procedure; these are discussed in Appendix C. Table 6 reports estimated lower- and upper-bound values of $\widetilde{r}_H$ for the three states, calculated separately for dropouts and graduates, as well as for the combined sample. The columns headed $\widetilde{r}_{.05}$ and $\widetilde{r}_{.95}$ provide the 5^{th} and 95^{th} quantiles of each bootstrap distribution, respectively; these quantiles are the limits of a 90% confidence interval. The columns headed "Point" provide the point estimate based on the actual data.

All of the lower-bound point estimates for $\widetilde{r}_H$ are small, with the largest being 0.066 for Illinois dropouts. For the overall sample—which includes both dropouts and graduates, appropriately weighted by their share of the constant-population effect—the largest lower-bound point estimate is 0.035, for Wisconsin. While these figures are close to zero, the confidence intervals show that they are nonetheless significantly different from 0. Moreover, the confidence intervals for the lower-bound estimates of $\widetilde{r}_H$ are generally quite tight; only the Illinois dropouts' 95^{th} quantile estimate exceeds 0.10.

Results are more varied for the upper-bound point estimates. On the low end, the Wisconsin overall point estimate is 0.054. On the other extreme, the upper-bound point estimate of 0.365 implies that optimal welfare benefits for Mississippi would be more than one-third greater if not for welfare migration. It is ironic to observe that welfare migration is likely to have a bigger relative impact on optimal benefits (0.095) in Illinois—the classic source of welfare migrants—than in Wisconsin—the classic recipient.

In evaluating these estimates, some caveats should be kept in mind. First, a given optimal benefits ratio implies different absolute benefit effects for different states, since the level of benefit

note Appendix A.2 on page 29; see Appendix C for details.

generosity differs so much. Second, I make no claim that actual state benefits would rise by the estimated amount if welfare migration were eliminated (even if we knew that, say, the upper-bound estimates were correct). This would be true only if state policymakers actually understand the role of marginal analysis as developed above and choose benefits accordingly. Third, while the discussion above assumes log utility for simplicity, underlying policymakers' preferences as reflected by ρ would make a substantial difference in translating the marginal cost ratio into the optimal benefits ratio.

As noted above, these estimates ignore in-migration effects. But even doubling the estimates reported in Table 6 would not significantly change my subtantive conclusions for the lower-bound estimates. For the upper-bound estimates, in-migration could be an important factor, though moreso in lower-benefit states. A more definitive analysis of the importance of in-migration would require estimating welfare's effects on in-migration, which is beyond the scope of this paper.

One final issue of interest concerns the importance of incorporating the lifecycle when thinking about welfare migration. The probit estimates imply that welfare migration effects operate early in the lifecycle for 1980 never-married dropouts, and these early effects will persist throughout the lifecycle. This is important because conditional on welfare participation, welfare income will be greater later in the lifecycle, when women have more children. Thus the effect of benefit levels on migration depends importantly on the profile of effects over the lifecycle. I illustrate the importance of the profile in Figure 1 by graphing the simulated impact (based on characteristics for Illinois and the sample of never-married dropouts) of a 10% benefit cut on one-year out-migration rates and on the "holdover rate", i.e., the probability that a woman has yet to leave the state, by oldest child's age. This figure shows that the impact of the benefit cut on the one-year migration rate is relatively great when oldest children are young. It then tapers off quickly, reaching 0 by the time oldest child's age is about 8. By contrast, the impact on the holdover rate increases until oldest child's age equals 6, where it reaches a maximum of nearly three percentage points. The holdover rate remains above two percentage points throughout the 12-year lifecycle, even though the one-year migration impacts can safely be ignored for mothers of older children. This illustrates the importance of accounting for the lifecycle when welfare migration varies over it.

4.3 Evaluating recent research on welfare migration

We can now investigate the conditions under which alternative measures of welfare migration appropriately capture the impact of welfare migration on the optimal benefits ratio. The discussion in Appendix B implies that we may write $\widetilde{r}_H = \sum_a \theta_a \widetilde{r}_{H,a}$, where θ_a is the share of the total constant-population effect accounted for by women whose oldest child is age a and $\widetilde{r}_{H,a}$ has the obvious definition. Neither Levine & Zimmerman (1999) nor Meyer (1998) address lifecycle issue, so in mapping their results into the previous subsection's theoretical framework I will assume that θ_a and $\widetilde{r}_{H,a}$ are uncorrelated.[22] Since the weights sum to unity and the lifecycle is 12 years long, this assumption implies $\widetilde{r}_H = (1/12)\bar{\widetilde{r}}_H$, where $\bar{\widetilde{r}}_H$ is the simple mean of age-specific $\widetilde{r}_{H,a}$ values. As a consequence, one can focus on average values of S, N, S', and N'.

LZ focus on the percentage effect of welfare migration on the steady-state welfare caseload. Because their preferred point estimates actually are wrong-signed from the point of view of welfare migration, LZ focus on the 95$^{\text{th}}$ quantile of the sampling distribution for their estimate, which yields a 4% impact on the steady-state caseload. LZ calculate this effect based on a specification focusing on dummy variables indicating whether the largest available increase in a woman's benefit via migration is at least \$100. Imagine an Illinois policymaker who is considering reducing benefits so that no resident of any other state can gain at least \$100 by moving in. Based on my 1980 data, Illinois would have had to reduce its two-person maximum benefit \$559 to \$409 to do so, since the lowest benefit was Mississippi's \$310. This \$150 difference represents about a 30% reduction, while my simulation concerned a 10% cut, so I scale down LZ's estimated percent-of-caseload effect by a factor of three.

To map the percent-of-caseload effect into my framework, I multiply and divide the last term in (7) by P, yielding $\widetilde{r}_{LZ} = (\Delta N/P)(P/N)(S/\Delta S)$. Here $\Delta N/P$ is the percent-of-caseload effect. In their calculation, LZ assume that welfare migrants receive the maximum benefit for three years but have no later welfare income. I thus assume that S equals the maximum monthly welfare-plus-Medicaid benefit for three years. Using my data, this benefit for an Illinois woman with the fertility profile assumed in my simulation works out to \$1,224 per month, or \$44,064 over a three-year period. The constant-population effect that results from the benefit cut occurs in all 12 years of the lifecycle and, based on my simulation results, amounts to \$11,232. The remaining term to fill in is the welfare participation rate P/N; among single mothers in my sample, this rate is 0.28 in my 1980 sample and 0.24 in my 1990 sample, so I use 0.26.

Putting all of these terms together thus yields $\widehat{\widetilde{r}}_{LZ} = (0.04/3) \times 0.26 \times (44064/11232) = 0.014$ as

[22] As figure 1 indicates, doing so will tend to understate the extent of welfare migration; for comparison, however, this is the right way to consider results in LZ and Meyer.

a rough estimate of the effect on the Illinois optimal benefits ratio based on LZ's 95% upper-bound estimate. This is somewhat smaller than my lower-bound Illinois point estimate and several times smaller than the 95^{th} quantile for my Illinois lower-bound estimate, as well as the upper-bound point and 95^{th} quantile estimates. This analysis shows that my point estimates are clearly much greater than even the 95^{th} quantile of LZ's. Because LZ use panel data on one-year migration, it is not possible for me to assess directly the source of these differences. However, a likely explanation is LZ's inclusion of repeated observations on all women in their NLSY sample. This research design likely biases estimates heavily toward finding no statistically significant effect, since it is inappropriate to treat location as reassigned in an *iid* fashion each year.

Next consider the measure used by Meyer (1998), the migration-induced increase in population in states with above-median benefits. Focusing only on highschool dropouts, Meyer estimates that the population share of 1980 gross flows out of below-median states are 0.00039 to 0.00714. This may be thought of as an estimated range for $\Delta N/N$. These impact estimates are based on the difference in migration rates across states with below- and above-median benefits. Based on my data, equalizing 1980 maximum benefits with a randomly chosen below-median state would require a randomly chosen above-median state to cut its benefit by about 30 percent. To obtain an estimate of $\widetilde{r}_M$, this interval will have to be scaled by $S/\Delta S$. Since my simulation-based estimates of S and ΔS are based on only a 10 percent cut, I divide Meyer's estimates by 3, which yields $[0.00013, 0.00238]$ as the range of adjusted estimates for the percentage increase in population. Considering only dropouts in Illinois yields an estimate of ΔS equal to \$186 per month, and as above my upper-bound monthly value of S is \$1,224. Thus we have an upper-bound estimate of $S/\Delta S = 1224/186 = 6.58$. Scaling the percentage increase in population interval by this factor yields an interval for $\widetilde{r}_M$ of $[0.001, 0.016]$. Both of these figures are much smaller than my estimates—below the corresponding 5^{th} quantile estimate for Illinois dropouts in Table 6. If correct, these estimates would certainly justify Meyer's conclusion that "state governors and legislators should be more worried about the effect of benefit levels on participation by their own constituents, than about the effects of benefits on migration of single mothers" (p. 25).

However, the discussion above of Figure 1 suggests that ignoring the lifecycle, as Meyer does, will tend to greatly understate the importance of welfare migration. To see whether the large difference in Meyer's and my estimated $\widetilde{r}_H$ values are due to this fact or something else, I estimated models like those in Table 3, but without the oldest child's age interaction term. Such models can be thought of as roughly comparable to Meyer's measure of flows out of lower-benefit states, provided that I evaluate the marginal effects for an appropriately sized change in the maximum benefit level. When I weight the results for never-married dropouts (which are substantial) and ever-married dropouts (which are essentially 0) by their population shares, the resulting estimates are about 30% larger than Meyer's. This is not enough to explain the full difference in estimated $\widetilde{r}_H$ values. The difference is thus partly due to different underlying estimates, and partly due to my incorporation of lifecycle effects in this paper.

5 Welfare migration and cross-sectional studies

Many researchers use cross-sectional variation in welfare benefit generosity to identify incentive effects of welfare benefit generosity on a wide array of outcomes. But as Moffitt (1992, p. 36) writes, "if residential location is endogenous, cross-state variation in benefits is as well and therefore the studies that use such variation may yield biased and inconsistent parameter estimates." My estimates appear to suggest substantial endogeneity of location relative to total migration for 1980 never-married dropouts, the most frequent users of welfare in my data, so this issue should be addressed. A very straightforward approach is to ask whether including variables related to locational choice on the right-hand side affects the estimated impact of welfare benefit generosity on welfare income.

The first column of Table 7 reports estimated CLAD coefficients for the age-0 and oldest child's age interaction effects of the maximum welfare benefit; these point estimates come from the model used to predict never-married dropouts' welfare income for the simulation above. Standard errors for these estimates were computed using 250 bootstrap replications and are discussed in section Appendix C.2. The estimates show that an increase of \$100 in the monthly maximum two-person benefit is associated with a \$151 increase in the latent index for monthly public assistance income, an effect that is essentially invariant with respect to oldest child's age.[23] The next column presents estimates computed with the addition of four migration-related variables: a dummy variable indicating whether the woman moved in the five-year period before the Census, a continuous variable

[23]The fact that the estimate exceeds \$100 is partly due to the use of the two-person benefit in a sample whose members frequently have more than one child. However, it is probably also true that the single-index model is mis-specified in the sense of White (1982) and Kim & White (forthcoming). Substantively, an estimate above \$100 suggests that the effect on the intensive margin may exceed the extensive effect. The estimates reported here are thus best viewed as a mixture of these two effects in a heterogeneous population.

equal to the amount by which the state maximum benefit for which she was eligible changed if she moved (and 0 otherwise), and the interaction of each of these variables with oldest child's age. The estimated coefficients for these four variables are all statistically insignificant. Moreover, adding them to the model makes essentially no difference in the estimates of the maximum benefit coefficients, as shown in the table's third column. It is striking that there is so little evidence that migration affects these coefficient estimates, given that never-married dropouts are the most benefits-responsive women in my migration results. Moreover, if endogeneity were likely a problem, one would certainly expect to see it when the dependent variable is welfare income.

However, even in cases when migration is systematically associated with the dependent variable, there will likely be little reason for concern over endogeneity. Welfare migration is in some cases large compared to the total number of *migrants*, but what matters for endogeneity bias is how large welfare migration is compared to the total *population* in the estimated model. To see why, consider a stripped-down model for an arbitrary, uncensored dependent variable: $y_{is} = \alpha + \beta b_s + u_{is} + \epsilon_i$, where ϵ_i is independent of b_s and the only new element is the unobserved term u_{is}, which represents heterogeneity due to the fact that person i chooses state of residence s.

Suppose for simplicity that there are only two possible locations, one with high benefits (denoted h) and one with low (denoted l); that the pre-migration populations in the two states are each of size 1 and have identical mean-0 distributions of u; and that the only movers are welfare migrants from l to h, who have mass m. The OLS estimate of β is the cross-state difference in means, $\overline{y}_h - \overline{y}_l$, with expectation $\beta + E[u_h] - E[u_l]$. After migration has occurred, $E[u_l] = E[u|\text{non-wm}] = -[m/(1-m)]E[u|\text{wm}] < 0$ (where "wm" and "non-wm" refer to welfare migrants and non-welfare migrants), and $E[u_h] = (1+m)^{-1}\{1 \times 0 + mE[u|\text{wm}]\} => 0$, so that the bias in the OLS estimate of β equals $2m/(1-m^2)]E[u|\text{wm}]$. Even with $m = .02$, which would be large, the first term is only 0.04. Thus a relative bias of, say, 20 percent would require that $E[u|\text{wm}]$, the heterogeneity among welfare migrants, be five times the true effect, β. Heterogeneity of this size would imply that $E[u|\text{wm}]$ in Table 7 is \$750 per month in welfare income. This might be possible relative to all single mothers, but the population in the reported CLAD models consists of only never-married dropouts, who already have mean public assistance income equal to \$410 per month. An increment of \$750 would essentially mean that all welfare migrants had moved to the very highest-benefit state in the U.S. and collected the maximum benefit for a family of size four. Thus even though *all* migration in this example is welfare migration, the bias cannot be very large. It is thus clear that welfare migration is unlikely to create serious bias when using cross-state variation in welfare policies.

6 Conclusion

In this paper, I find substantial evidence of welfare migration for 1980, with further evidence that welfare migration varies over the lifecycle, especially among never-married dropouts, as theory predicts. Results for 1990 data are mixed and difficult to interpret. An important advance made here is the development of a theory-driven framework for ascertaining whether welfare migration has important effects on optimal benefit levels. For a higher-benefit state like Wisconsin, the answer appears to be no. Unmeasured in-migration effects could raise this somewhat, but even a doubling of this effect would constitute only a moderate impact. For lower- and middle-benefit states like Mississippi and Illinois, there is a greater possibility of large relative effects. I thus conclude that the (relative) bottom may be considerably lower in lower-benefit states than in higher-benefit ones. At the same time, there is no evidence that failing to account for welfare migration leads to noticeably different estimates of the impact of welfare benefit generosity on welfare income. The key factor here is the magnitude of welfare migration relative to the population (which is large), not to the population of migrants (which is small).

One further issue bearing discussion is welfare reform. The data used here all pre-date the large wave of state and federal welfare reforms that began in the early 1990s, so the estimates are not themselves subject to concerns regarding these reforms. However, one might ask whether the topic of welfare migration remains relevant, now that federal law stipulates a five-year lifetime limit on receipt of cash assitance, especially since cash welfare is no longer an entitlement. Nonetheless, a reasonable case can be made that welfare migration remains at least as important a topic as previously. This is partly true because it is always possible that at some future date, policy cycles will bring back interest in programs without time limits and other recent innovations.

Moreover, it is difficult to overstate the potential importance of welfare migration to the viability of a decentralized welfare system like the one implemented in PRWORA. While states were always allowed to set their own benefit levels under AFDC, they now have much more discretion regarding work incentives, time limits, sanctions, work requirements, and other program features. One likely response to welfare migration is for states to either cut back on costly carrots or increase the intensity of sticks. To put it simply, the race to the bottom has many more lanes after PRWORA. Whether my estimates of $\widetilde{r}_H$ carry over to qualitatively new features of the post-PRWORA welfare system is a remaining empirical question.[24]

As for time limits, states can continue paying benefits as long as they do not use federal funds for payments after the time limit has expired.[25] Federal TANF dollars are fungible in the sense that states could choose to assist time-limited family A with "state" funds and non-time-limited family B with "TANF" funds. Hence PRWORA does not effectively *prevent* states from providing long-term assistance. Rather, it *allows* them to *choose* not to, again opening the possibility of a race to the bottom. Also, since states generally have not coordinated their records of welfare receipt, a long-term recipient living in a state that enforces a binding lifetime limit on assistance faces a maximum benefit of 0 in that state. If her state's welfare agency does not coordinate its records with other states, she may be able to extend her eligibility by up to five more years via migration. Hence, migration may be the only way for time-limited mothers to retain *de facto* eligibility.

Lastly, as a result of the major changes in welfare policy since the mid-1990s, transparent variation in locational incentives is likely to be increasingly hard to come by. Many different aspects of state safety nets—child care subsidies, time limits, family caps, sanctions, positive work incentives, job training, job placement—have changed at the same time. Since each state has its own bundle of reform initiatives, the prospects are poor for using cross-state comparisons to identify the effects of any one reform. The relatively stable structure of AFDC rules between the mid-1970s and 1990 thus provides arguably the best data for use in welfare migration studies.

[24] De Jong & Graefe (2002) use panel data from the Survey of Income and Program Participation, together with measures of the stringence of post-PRWORA state welfare programs, and conclude that more stringent reforms are associated with migration in the direction predicted by theory. Kaestner, Kaushal & Ryzin (2001) use Current Population Survey data and dummy variables indicating specific state policies, finding that reforms are associated with increased mobility. Much of this mobility appears to be within-state, raising questions about whether women are "shopping around" for better welfare packages or simply looking for better employment opportunities.

[25] Grogger & Michalopoulos (2003) note that some states have chosen to use this option.

Appendix A Data

Appendix A.1 Sample construction

The microdata used here come from the 5% Public Use Microdata Samples (PUMS) of the 1980 and 1990 U.S. Censuses. Each Census provides data on state-to-state migration over the five-year period preceding the Census. I do not use 1970 data because the Supreme Court struck down residency requirements between 1965 and 1970, and because sample sizes in the 1970 PUMS would be substantially smaller. The samples are conditioned on a number of selection criteria. First, I select only women who reported being the head of their household at Census time. Second, I select only women for whom the reported number of children ever born matches the number listed as own-children of the household head. Third, I select only those women whose oldest child had been born five years before the Census and was 16 or younger. For women who had yet to give birth five years before the Census, oldest child's age could in principle be endogenous to location. I drop women whose oldest child was 17 at Census time because some 17-year-olds may have left home. Fourth, I drop women who reported a disability. I do so because the PUMS provides only a single aggregate variable for public assistance income (which I use for the simulation below), and disability payments (which are set at the national level) are a primary alternative source of public assistance income. Fifth, I discard observations with allocated data for variables used in this paper. Summary statistics for a number of demographic variables, as well as the state-level variables described below, are provided in Appendix Table 8.

Appendix A.2 Construction of state-level variables

Maximum welfare benefits are calculated as follows. For a given year, I first take each state's maximum welfare payment for a family of two. I then multiply by 0.7, because the Food Stamps program has traditionally imposed an implicit 30% tax on AFDC income. I then add the maximum Food Stamps payment for a family of two, because AFDC eligibility brings categorical eligibility for Food Stamps, and virtually all AFDC participants receive Food Stamps. To account for inflation, I use the Consumer Price Index (CPI), with all monetary figures in the paper provided in 1997 dollars.[26,27] Lastly, to avoid year-specific noise, all state-level variables are averaged over a five-year period centered on the Census year. For example, 1980 state-level variables' are constructed as the simple average value of those variables for the period 1978–82.

I obtained data on average Medicaid expenditures for a family of four from Robert Moffitt's website (`http://www.econ.jhu.edu/people/moffitt/DataSets.hmtl`). This variable is documented in detail there.

Data on school expenditures and school enrollment come from the *Digest of Education Statistics* for 1976 (Table 31), 1977/78 (Table 70), 1993 (Table 43), and 1995 (Table 159).

Data on state income used to construct *per capita* state income come from the Bureau of Economic Analysis's Regional Economic Information System State Annual Summary Tables (SA1-3, SA51-52), July 1999 release. Population data come from the *Statistical Abstract of the United States* for 1982 and 1983 (Table 11) and 1988 (Table 21), as well as *Public Education Finances: 1990-91*, published by the US Commerce Department Economics and Statistics Administration (Table 22; source for the population data "represents resident U.S. population from the 1990 Decennial Census of Population and Housing").

The marriageable men ratio was constructed following the methodology suggested by Wood (1995). I use PUMS data to construct a race- and state-specific marriageable-men index. I first count the number of men of a given race (White, Black, or Other) in a state with 1979 wage and salary income of at least $8,000 in nominal dollars (roughly the poverty line for a family of four); for the 1990 PUMS, I use a $13,300 nominal cutoff. The total number of women—married or not—of the given race is then counted and used as the denominator of the marriageable-men ratio. To avoid small sample sizes, I do not condition on age groups, instead including all men and women aged 18–54 in these counts. There are at least 20 men and at least 20 women in each cell with the exception of Blacks in South Dakota for the 1980 Census. I drop 1980 South Dakota observations from the estimation.

[26] I use a family of two for simplicity. One could try to use the additional variation in incremental state-level benefits provided as family size changes. However, there is no obvious way to account for fertility expectations, and in any case incremental benefits are typically much less than the benefit for a family of two.

[27] I am thus assuming that Food Stamps are worth face value. Recent evidence presented in Whitmore (2002) suggests that Food Stamps are an inframarginal subsidy for most recipients, with the rest valuing them at 80% of face value on average; this finding is consistent with other work on the cash value of Food Stamps.

Appendix A.3 Educational subgroups

For the 1980 PUMS, I code women as dropouts if they have completed fewer than 12 years of education. I code highschool graduates are women who have completed at least 12, but fewer than 16, years of education. I code college graduates as those who have completed at least 16 years.

The 1990 PUMS reports educational attainment partly by degree and partly by years completed. I code 1990 PUMS women as dropouts if they report having completed education less than or equal to 12th grade, but do not report being/having a "High school graduate, diploma or GED". I code highschool graduates as those who report at least this much education, but not as much as a Bachelor's degree. I code college graduates as women with at least a Bachelor's.

Appendix B A framework for assessing welfare migration

The first-order condition with respect to welfare benefits is to set $b^* = [\lambda^* C'(b^*)]^{-1/\rho}$, where asterisks denote optimality. Now consider the counterfactual world in which there is no welfare migration, with $\widetilde{C}$ replacing C as the welfare cost function. Letting two asterisks denote optimality in the no-welfare-migration world, the actual and counterfactual first-order conditions are related by

$$R \equiv \frac{b^{**}}{b^*} = \left(\frac{\lambda^*}{\lambda^{**}} \frac{C'(b^*)}{\widetilde{C}'(b^{**})} \right)^{\frac{1}{\rho}}, \tag{4}$$

which is the ratio of optimal benefits in the no-welfare migration world to optimal benefits given the observed level of welfare migration. The first term in parentheses on the right-hand side of (3) is the ratio of the multipliers, which can be interpreted as the relative real-world shadow price of government revenues. I will take this relative price to be unity, which is reasonable because welfare expenditures are a very small fraction of total state expenditures. For example, consider state expenditures that occur because of welfare policies. At its peak in 1994, total state spending on AFDC was \$13.9 billion (U.S. House of Representatives, Committee on Ways and Means (2000, p. 404)). AFDC brings with it categorical eligibility for Medicaid, and for 1994, a total of about \$31 billion was spent on Medicaid for AFDC recipients (see U.S. Bureau of the Census (1997, Table 168)). Of this sum, less than half was paid for by states (since the federal matching rate for Medicaid expenditures is at least 50% for every state). Thus total state spending on AFDC and related Medicaid spending was less than \$30 billion for 1994. Total state government expenditures in that year were approximately \$775 billion (see U.S. Bureau of the Census (1997, Table 477)), so AFDC payments accounted for less than 4% of state expenditures.[28]

The second term in parentheses on the right-hand side of (3) is the ratio of the marginal cost of welfare generosity given observed welfare migration to the marginal cost in the no-welfare-migration world. As stated in the text, I will assume $\rho = 1$ (i.e., log utilities) for the rest of the paper.

Ignoring lifecycle considerations for the moment, let the number of single mothers living in a state with benefit b be given by $N(b)$, and suppose for the moment that all of them receive the same welfare payment, denoted $S(b)$. We may write the state's expected welfare expenditures as

$$C(b) \equiv S(b)N(b), \tag{5}$$

and the marginal cost of benefits as

$$C' \equiv SN' + S'N. \tag{6}$$

Since by definition $N' = 0$ in the absence of welfare migration, the constant-population effect is the entire marginal cost of benefit generosity in the no-welfare-migration world, i.e., $\widetilde{C}'(b) = S'(b)N(b)$. Thus the optimal benefits ratio may be written

$$R = \frac{S(b^*)N'(b^*) + S'(b^*)N(b^*)}{S'(b^{**})N(b^{**})}.$$

The calculation of the marginal cost ratio will be greatly simplified if we can take the constant-population effect to be roughly equal in the observed and no-welfare-migration worlds, so that $S'(b^*)N(b^*)$ is roughly equal to $S'(b^{**})N(b^{**})$.[29] Under this assumption we may focus on estimating

[28] It may not be obvious why the multipliers will be essentially unchanged when potentially large relative changes occur in an expenditure category that is small compared to total expenditures. The first-order condition for optimality is that the multiplier equal the marginal social benefit of expenditures per marginal dollar spent on each category, i.e., $\lambda^* = F_{z_k}/Q_{z_k}$ for the k^{th} publicly provided good. Assume that there is at least one expenditure—say, roads construction—for which welfare migration does not alter the marginal cost Q_{z_k}. Now, (*i*) relatively little is spent in total on welfare, only some of which would be affected by the absence of welfare migration, (*ii*) any savings or losses to a state from the elimination of welfare migration will be spread over multiple uses, and (*iii*) roads construction is expensive, so small changes in spending on z_k are unlikely to increase its quantity much. Unless the state is virtually roadless, so that the marginal social benefit of a very small increase in roads is enormous, it is reasonable to believe that in general equilibrium, the marginal social benefit F_{z_k} of roads construction will be approximately unaltered by any increased spending that results from eliminating welfare migration. With both marginal cost and marginal social benefit roughtly constant for this one publicly provided good, the multiplier must also be roughly constant.

[29] Formally, consider a Taylor series expansion of $\widetilde{C}'(b^{**})$ around b^*, which yields $S'(b^*)N(b^*) + (b^{**} -$

$$\widetilde{R} \equiv \frac{S(b^*)N'(b^*) + S'(b^*)N(b^*)}{S'(b^*)N(b^*)} = 1 + \frac{S(b^*)N'(b^*)}{S'(b^*)N(b^*)}, \tag{7}$$

as in the main text.

It remains to incorporate lifecycle issues into this framework, as well as to account for the fact that I do not have estimates of the impact of welfare benefit levels on in-migration. Suppose we consider only one cohort of women, defined by the year in which they have their first birth. For any oldest child's age a, let I_a be the number of women who move into the state when their oldest child is a, and let O_a be the number who move out. The total population of these women who live in the state when their oldest child is a is $N_a = H_a + I_a$, where $H_a \equiv N_{a-1} - O_a$ is the "holdover population" from the year when oldest child was $a-1$. Total spending is $\sum_a [S_{Ha}H_a + S_{Ia}I_a]$, where subscripts denote the segment of population on which spending is done, and we may use (7) to write $\widetilde{r}$ as $[\sum_a (S'_{Ha}H_a + S'_{Ia}I_a)]^{-1}(S_{Ha}H'_a + S_{Ia}I'_a)$, where S_H and S_I are, respectively, per capita welfare expenditures among holdovers and in-migrants. Defining $\theta \equiv \sum_a (S'_{Ha}H_a) / \sum_a [S'_{Ha}H_a + S'_{Ia}I_a]$,

$$\widetilde{r} \equiv \theta\widetilde{r}_H + (1-\theta)\widetilde{r}_I, \tag{8}$$

where $\widetilde{r}_H \equiv S_H H'/S'_H H$, and similarly for $\widetilde{r}_I$. The optimal benefits ratio is a weighted average of the ratios for holdovers and for in-migrants, where the weights are the fraction of the constant-population effect accounted for by each group. Now, $\widetilde{r}_H$ will yield an appropriate estimate of $\widetilde{r}$ if either θ is close to unity or $\widetilde{r}_H$ and $\widetilde{r}_I$ are similar. If we assume that $S'_H \approx S'_I$, then the weights on each $\widetilde{r}$ term are approximately equal to the holdover and in-migrant population shares. Total migration in my sample is less than 10%, so θ likely exceeds 0.9. Ignoring in-migrants thus is problematic only if $\widetilde{r}_I$ is much larger than $\widetilde{r}_H$. As an example, suppose that $\widetilde{r}_I$ is large—say, 10 times the size of $\widetilde{r}_H$, and suppose the constant-population effect is proportionate to population, so that the greater value of $\widetilde{r}_I$ is caused by relative differences in the welfare migration effect. Then we can again take $\theta = 0.9$, so $\widetilde{r} = 10\,\widetilde{r}_H \times 0.1 + \widetilde{r}_H \times 0.9 = 1.9\,\widetilde{r}_H$. Thus when $\widetilde{r}_H$ is small enough, we can safely ignore in-migration.

$b^*)[S''(b^\dagger)N(b^{**}) + S'(b^*)N'(b^\dagger)] + (b^{**} - b^*)^2 N'(b^\dagger)S''(b^\dagger)$ for some $b^\dagger \in [b^*, b^{**}]$. The leading term will dominate for intervals $[b^*, b^{**}]$ such that either $b^{**} - b^*$ is relatively small, or neither S' nor N is too elastic, so that bS'' is small relative to S' and bN' is small relative to N. Now, S'' will be small relative to S' as long as first-order effects of the benefit level on expected spending are large compared to higher-order effects. N' will be small relative to N as long as welfare-induced migration is not too large relative to the overall population. This is surely an appropriate assumption, since *total* five-year state-to-state migration is under 10% in my sample.

Now relax the assumption that all women receive identical welfare payments. Consider the impact, among holdover women whose oldest child is a, of a cut in the maximum benefit equal to Δb. This impact is

$$\Delta C_a = S_{a,\Delta H_a}(b)\Delta H_a + \Delta S_{a,H_a(b-\Delta b)} H_a(b-\Delta b). \tag{9}$$

The first product on the right-hand side is the discrete-change analogue to the welfare migration effect $S_H H'$, with ΔH_a being the change that occurs in the size of the holdover population having oldest child's age a when the maximum benefit is cut by Δb. The spending term $S_{a,\Delta H_a}(b)$ is the amount of spending that would have been done, in the absence of a benefit cut, on oldest child's age-a women who are caused by the benefit cut to move out of the state. The second product on the right-hand side is the discrete-change analogue to the constant-population effect $S'_H H$, with $H_a(b-\Delta b)$ being the holdover population now that the maximum benefit equals $b-\Delta b$. The change in spending for these women given the size of the benefit reduction is $\Delta S_{a,H_a(b-\Delta b)}$. Because my simulation uses a 10% benefit cut, equation (9) provides the correct theoretical construct. The total impact of the benefit cut over the lifecycle is $\sum_a \Delta C_a$, with the respective welfare migration and constant-population effects being $\sum_a S_{a,\Delta H_a}(b)\Delta H_a$ and $\sum_a \Delta S_{a,H_a(b-\Delta b)} H_a(b-\Delta b)$, respectively. Thus in constructing the welfare migration effect, we weight the oldest child's age-specific change in the holdover population by the expenditures that would be made on welfare migrants if there were no benefit cut. Similarly, for the constant-population effect we weight the holdover population that will exist if there is a benefit cut by the change in spending on this population given the cut. The estimate of $\widetilde{r}_H$ is then simply the ratio of these simulated effects.

Appendix C Mathematical details of the simulation procedure

Let the set of possible marital history and status states be $\mathcal{M} \equiv \{\text{never, ever-and-single, ever-and-married}\}$, let the set of possible oldest child's ages be $\mathcal{A} \equiv \{0, 1, 2, \dots, 11\}$, and let the set of educational attainment states be $\mathcal{D} \equiv \{0, 1\}$ (indicating whether or not the woman is a dropout). Let the set of state-level characteristics be $\{X_s\}$ for $s \in \mathcal{S} \equiv \{\text{Mississippi, Illinois, Wisconsin}\}$. Let $\mathcal{F} = \{0.9, 1\}$ be the set of benefit scale factors with elements denoted by f; if $f = 1$, then a state's actual benefit value will be used, and if $f = 0.9$ then 90% of that value will be used. For any given woman, the set of possible states of the world I will consider is the Cartesian product $\Omega \equiv \mathcal{M} \times \mathcal{A} \times \mathcal{D} \times \mathcal{S} \times \mathcal{F}$, with elements of this set denoted ω. For convenience, define the marital history/status implied by ω as $m(\omega)$, the educational attainment as $d(\omega)$, the oldest child's age as $a(\omega)$, and the state characteristics as $s(\omega)$. Let $y_i(\omega)$ be (true, counterfactual) expected spending on woman i if she is in state ω. Also let $p_i(\omega) = p_i(m(\omega), d(\omega)) \equiv Pr[i \text{ is in state } \omega | a(\omega), d(\omega), s(\omega)]$ be the probability of having marital history/status and educational attainment $(m(\omega), d(\omega))$ given the oldest child's age, dropout status, and state-level characteristics. I take dropout status to be fixed over the life-cycle, so $p_i(\omega) = 0$ whenever woman i's actual dropout status $d_i \neq d(\omega)$. Expected counterfactual spending on woman i given ω and given that she remains in U.S. state $s(\omega)$ is thus $y_i(\omega)p_i(\omega)$. Lastly, let $X_i(\omega)$ be the list of all individual demographic variables that vary with ω.

As noted in the main text, expected spending given ω is predicted using CLAD coefficient estimates, with the prediction left-censored at 0. Given individual and state-level characteristics $X_{is(\omega)} \equiv (X_i(\omega), X_s(\omega))$, woman i's expected welfare income is estimated as $\widehat{y_i}(\omega) = \max[0, X_{is(\omega)}\hat{\beta}_{CLAD,md(\omega)}]$, where $md(\omega)$ denotes the marital history and dropout status implied by ω. For the lower-bound estimates, $\widehat{y_i}(\omega)$ is used in equations (10) and (11) below. For the upper-bound estimates, I instead use the maximum benefit level available given the number of children at the simulated oldest child's age. To account for state Medicaid expenditures, I take the predicted public assistance income and divide by its maximum possible value given the state maximum benefit level and the number of simulated children in the household (so this ratio is unity in the upper-bound case). I then multiply this ratio by the simulated household size, divided by four (I make this latter adjustment because I have data on Medicaid expenditures only for a family of four), and multiply by the average Medicaid expenditure variable in Moffitt's dataset (described in Appendix A.2 on page 29).

To account for outmigration, we need to multiply the predicted conditional-on-state of residence spending term $\widehat{y_i}(\omega)\widehat{p_i}(\omega)$ by the probability the woman has not moved out of state by the time her oldest child's age is $a(\omega)$. For the estimated probability of one-year migration when oldest child's age is a, I use $\widehat{m}_i(\omega) \equiv [1 - \Phi(X_{is(\omega)}\hat{\beta}_{3,md(\omega)})]^{1/5}$, where $\hat{\beta}_{3,md(\omega)}$ is the third-column estimate from Table 3 for a woman having marital history and educational attainment md, as specified by ω.[30] The probability that woman i remains in the given U.S. state when her oldest child is age a is

[30] I am thus assuming that one-year outmigration rates are constant over any five-year period. This assumption will not generally be exactly correct, since five-year migration rates involve overlapping years and are not constant. However, it is a reasonable rough approximation.

estimated as $\widehat{H}_{i,a(\omega)}(\omega) \equiv \widehat{H}_{i,a(\omega)-1}(\omega_{a-1})(1-\widehat{m}_i(\omega))$, where ω_{a-1} refers to ω with all components related to $a(\omega)$ altered to instead reflect oldest child's age $a(\omega)-1$ and $\widehat{H}_{i,-1} \equiv 1$. Total expected spending on woman i given benefit ratio f is thus $\widehat{c}_i(f) \equiv \sum_\omega (1-\widehat{H}_{i,a(\omega)}(\omega))\hat{y}_i(\omega)\widehat{p}_i(\omega)$, and total expected spending given f is $\widehat{C}(f) \equiv \sum_i \widehat{c}_i(f)$. The difference $\Delta\widehat{C} \equiv \widehat{C}(1) - \widehat{C}(0.9)$ is the savings achieved when the monthly maximum benefit is cut by 10%.

Estimating $\widetilde{r}_H$ requires separating $\Delta\widehat{C}$ into welfare migration and constant-population effects. Let $\omega_{f=f'}$ equal ω with the component f set to the particular value f'. Then the welfare migration effect is

$$\sum_i \sum_{\omega_f} \Delta\widehat{H}_{i,a(\omega)}(\omega)\widehat{y}_i(\omega_{f=1})\widehat{p}_i(\omega_{f=1}), \tag{10}$$

where $\Delta\widehat{H}_{i,a(\omega)}(\omega_{-f}) \equiv \widehat{H}_{i,a(\omega)}(\omega_{0.9}) - \widehat{H}_{i,a(\omega)}(\omega_1)$ is the impact on woman i's migration probability of the benefit cut when all characteristics except for f are held constant at ω values. The estimated constant-population effect is

$$\sum_i \sum_\omega (1-\widehat{H}_{i,a(\omega)}(\omega_{f=0.9}))[\widehat{y}_i(\omega_{f=0.9})\widehat{p}_i(\omega_{f=0.9}) - \widehat{y}_i(\omega_{f=1})\widehat{p}_i(\omega_{f=1})]. \tag{11}$$

Appendix C.1 Convergence issues in implementing the bootstrap

The statistic $\widetilde{r}_H$ is unlikely to be asymptotically pivotal (meaning that its asymptotic distribution probably depends on unknown parameters), so a large number of bootstrap replications is necessary to get accurate confidence intervals using the percentile method; see Davison & Hinkley (1997) for a discussion of this issue.

My bootstrap procedure involves estimating 1000 sets of three probits, which necessarily converge given the global concavity of the log-likelihood function, and three CLAD models, which necessarily converge only when all the parameters are well identified. On a Pentium XXX running Red Hat linux, typical replications of the CLAD models took 1–2 minutes for the high school graduates sample (and half as long for the smaller dropouts sample). Unfortunately, the CLAD coefficients appeared to be poorly identified over some parts of the sample space, with iterations for some of the bootstrap replication samples continuing even after hours. Even when run in multiple processes in parallel, these replications greatly raised computational costs. I thus placed a 10-minute limit on each replication for graduates (and 5 minutes for dropouts). As a result, I lose 59 replications for graduates and 13 for dropouts; 3 of these "bad" replication samples overlap for dropouts and graduates, so for the overall sample I lose 69 replications. In constructing confidence intervals, it is unclear how to handle these "missing" coefficient vectors. The most conservative approach is to construct the lower confidence limit as if the estimated value of $\widetilde{r}_H$ for all bad replications would be below the minimum observed in the good replications, and analogously for the upper confidence limit. This is acceptable for the dropouts, since fewer than 50 replications were censored by the time limit. Unfortunately, given the larger number of bad replications in the high school graduates sample, conservative 90% confidence intervals are not identified this way. The most sanguine—and least conservative—assumption is that the event of hitting the 10-minute time limit is statistically independent of the sampling distribution for the estimates; under this assumption, I can simply use the percentile method based on the replications that converged, which is what I report in Table 6. More detailed statistics on the bootstrap distribution for my $\widetilde{r}_H$ estimates are available on request.

Appendix C.2 Standard errors for the CLAD estimates

For the model without migration variables in Table 7, 1000 replications are already available since I used them in the the simulation, but the model with migration variables has to be bootstrapped separately. Asymptotic normality of these two CLAD estimators allows use of standard errors rather than the percentile method. Since bootstrap standard errors are estimated as the sample standard deviation of the bootstrap distribution, and since the standard deviation is a continuous function of a sample moment, convergence occurs more quickly than for the percentile method. For this reason I use only 250 bootstrap replications of the models reported in Table 7. As discussed in Appendix Appendix C.1 on the preceding page, the CLAD estimator failed to converge quickly for some (15 in the first column and 14 in the second) of these 250 replication samples. The standard errors in Table 7 were computed under the assumption that this event is independent of the underlying estimate. One additional replication was dropped because it yielded an extreme outlier (more than 15 median units from the median) for the maximum benefit coefficient in the second-column specification.

References

Akerlof, G. (1978), 'The economics of 'tagging' as applied to the optimal income tax, welfare programs and manpower planning', *The American Economic Review* **68**(1), 18–19.

Baicker, K. (2001), 'Government decision-making and the incidence of federal mandates', *Journal of public economics* **82**, 147–94.

Bane, M. J. & Ellwood, D. T. (1994), *Welfare Realities: From rhetoric to reform*, Harvard University Press, Cambridge and London.

Blank, R. (1988), 'The effect of welfare and wage levels on the location decisions of female-headed households', *Journal of Urban Economics* **24**(2), 186–211.

Brown, C. C. & Oates, W. E. (1987), 'Assistance to the poor in a federal system', *Journal of Public Economics* **32**, 307–30.

Case, A., Hines, J. & Rosen, H. (1993), 'Budget spillovers and fiscal policy interdependence: Evidence from the states', *Journal of Public Economics* **52**(3), 285–307.

Cebula, R. J. (1979), 'A survey of the literature on the migration-impact of state and local government policies', *Public Finance* .

Davison, A. C. & Hinkley, D. (1997), *Bootstrap Methods and Their Application*, Cambridge University Press, Cambridge, United Kingdom.

De Jong, G. F. & Graefe, D. R. (2002), Welfare reform and migration: Moving to benefits; moving from restrictions, Working paper, Population Research Institute, The Pennsylvania State University.

Gelbach, J. B. (2000), The lifecycle welfare migration hypothesis: Evidence from the 1980 and 1990 censuses, Working Paper 00-01, University of Maryland.

Gelbach, J. B. (2002), Migration, the lifecycle, and state benefits: How low is the bottom? Typescript, available at http://glue.umd.edu/˜gelbach/papers/index.html.

Gramlich, E. M. & Laren, D. S. (1984), 'Migration and income redistribution responsibilities', *Journal Of Human Resources* **19**, 489–511.

Grogger, J. (2000), Time limits and welfare use, Working Paper 7709, NBER.

Grogger, J. & Michalopoulos, C. (2003), 'Welfare dynamics under time limits', *Journal of Political Economy* **111**(3), 530–54.

Hubbard, R. G., Skinner, J. & Zeldes, S. P. (1995), 'Precautionary saving and social insurance', *Journal of Political Economy* **103**(2), 360–398.

Kaestner, R., Kaushal, N. & Ryzin, G. V. (2001), Migration consequences of welfare reform, Working paper, NBER.

Kennan, J. & Walker, J. R. (2003), The effect of expected income on individual migration decisions, Working Paper 9585, NBER.

Kessler, D. P. & McClellan, M. B. (2000), 'Is hospital competition socially wasteful?', *Quarterly Journal of Economics* pp. 577–615.

Kim, T. & White, H. (forthcoming), 'Estimation, inference, and specification testing for possibly misspecified quantile regressions', *Advances in Econometrics* .

Kubik, J. D. & Meyer, B. D. (1992), Interstate migration of afdc recipients. cited in Meyer (1998).

Levine, P. B. & Zimmerman, D. J. (1999), 'An empirical analysis of the welfare magnet debate using the NLSY', *Journal of Population Economics* .

Maldonado v. Houstoun (1997), http://www.paed.uscourts.gov/documents/opinions/MALDONAD.HTM.

McKinnish, T. G. (2003), Evidence on cross-border welfare migration. Typescript, University of Colorado.

Meyer, B. D. (1998), Do the poor move to receive higher welfare benefits?, Working paper, JCPR.

Meyer, B. D. & Rosenbaum, D. (2001), 'Welfare, the earned income tax credit, and the laor supply of single mothers', *Quarterly Journal of Economics* **116**(3), 1063–1114.

Mirrlees, J. A. (1971), 'An exploration in the theory of optimum income taxation', *Review of Economic Studies* **38**(2), 175–208.

Moffitt, R. (1992), 'Incentive effects of the U.S. welfare system: a review', *Journal of Economic Literature* **XXX**, 1–61.

Moffitt, R., Ribar, D. & Wilhelm, M. (1998), 'The decline of welfare benefits in the U.S.: the role of wage inequality', *Journal of Public Economics* **68**, 421–52.

Pear, R. (1997), 'Judge rules states can't cut welfare for new residents', *The New York Times* p. A1.

Peterson, P. & Rom, M. (1989), 'American federalism, welfare policy, and residential choices', *American Political Science Review* **83**, 711–28.

Powell, J. (1984), 'Least absolute deviations estimation for the censored regression model', *Journal of Econometrics* **25**, 303–25.

Shapiro v. Thompson (1969). 394 U.S. 618.

U.S. Bureau of the Census (1997), *Statistical Abstract of the United States*, U.S. Government Printing Office, Washington.

U.S. House of Representatives, Committee on Ways and Means (2000), *2000 Green Book*, U.S. Government Printing Office, Washington.

Walker, J. R. (1994), Migration among low-income households: Helping the witch doctors reach consensus. Typescript.

White, H. (1982), 'Maximum likelihood estimation of misspecified models', *Econometrica* **50**(1), 1–25.

Whitmore, D. (2002), What are food stamps worth? Princeton University Industrial Relations Section, working paper #468.

Wood, R. G. (1995), 'Marriage rates and marriageable men', *The Journal of Human Resources* **30**(1), 163–93. Winter.

Wooldridge, J. M. (2002), *Econometric analysis of cross section and panel data*, The MIT Press, Cambridge, MA.

Ziliak, J. P. (2003), 'Income transfers and assets of the poor', *Review of Economics and Statistics* **LXXXV**(1), 63–76.

Table 1: Migration rates and the resulting change in maximum available benefit, 1975–80
(generating time: Thu Oct 2 17:00:35 2003 from file *change-in-benefits.ara*.)

			Movers only: change in maximum monthly benefit		
	Welfare	Any Move	Simple	Add Max. ben.	Add other controls
*A.*Treatment group					
Never-married dropouts	0.66***	4.648***	27*	25**	31***
	(0.01)	(0.435)	(15)	(11)	(11)
Max ben in 1975				-0.83***	-0.88***
				(0.01)	(0.01)
*B.*Relative to comparison group					
Ever-married dropouts	0.27***	-3.560***	54***	52***	55***
	(0.01)	(0.614)	(18)	(13)	(13)
Never-married HS grads	0.28***	-2.149***	25	17	17
	(0.01)	(0.607)	(19)	(14)	(14)
Ever-married HS grads	0.49***	-5.683***	43***	38***	41***
	(0.01)	(0.497)	(16)	(12)	(12)
Married Dropouts	0.62***	-3.164***	47***	60***	63***
	(0.01)	(0.457)	(16)	(11)	(12)
Married HS grads	0.65***	-5.928***	46***	42***	48***
	(0.01)	(0.442)	(15)	(11)	(11)
Max ben in 1975				0.83***	0.88***
				(0.01)	(0.01)
N	206,006	206,006	20,391	20,391	20,391
Other controls	No	No	No	No	Yes

Note: White-robust standard errors in parentheses.
***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

1. "Treatment" values report levels for never-married dropouts; in columns with controls, these levels are the estimated constant in a regression of the dependent variable on a constant and the de-meaned controls. Thus the reported value is the predicted migration rate or change in maximum benefits for a woman with mean characteristics.

2. "Relative to comparison group" figures are coefficients on dummy variables indicating the woman's marital history/educational attainment category, where the omitted category is never-married dropouts. Thus these figures are equivalent to adjusted values of the dependent variable for never-married dropouts, minus adjusted value for the relevant comparison group.

3. High school graduates may have some college, but have fewer than 16 years of education.

4. Women in sample must have given birth to the same number of children as the number of own-children who live in the household.

5. See discussion beginning on page 2 for details concerning controls.

Table 2: Migration rates and the resulting change in maximum available benefit, 1985–90
(generating time: Thu Oct 2 17:00:37 2003 from file *change-in-benefits.ara.*)

			Movers only: change in maximum monthly benefit		
	Welfare	Any Move	Simple	Add Max. ben.	Add other controls
*A.*Treatment group					
Never-married dropouts	0.58***	5.530***	12*	19***	25***
	(0.01)	(0.248)	(7)	(5)	(5)
Max ben in 1985				-0.90***	-0.91***
				(0.00)	(0.01)
*B.*Relative to comparison group					
Ever-married dropouts	0.22***	-2.919***	12	28***	32***
	(0.01)	(0.404)	(9)	(6)	(6)
Never-married HS grads	0.25***	-1.107***	13	14**	14**
	(0.01)	(0.314)	(8)	(6)	(6)
Ever-married HS grads	0.44***	-3.423***	17**	25***	30***
	(0.01)	(0.283)	(7)	(5)	(5)
Married Dropouts	0.54***	-1.748***	18**	29***	37***
	(0.01)	(0.272)	(7)	(5)	(5)
Married HS grads	0.57***	-3.276***	16**	22***	30***
	(0.01)	(0.253)	(7)	(5)	(5)
Max ben in 1985				0.90***	0.91***
				(0.00)	(0.01)
N	433,521	433,521	36,759	36,759	36,759
Other controls	No	No	No	No	Yes

Note: White-robust standard errors in parentheses.
***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

1. "Treatment" values report levels for never-married dropouts; in columns with controls, these levels are the estimated constant in a regression of the dependent variable on a constant and the de-meaned controls. Thus the reported value is the predicted migration rate or change in maximum benefits for a woman with mean characteristics.

2. "Relative to comparison group" figures are coefficients on dummy variables indicating the woman's marital history/educational attainment category, where the omitted category is never-married dropouts. Thus these figures are equivalent to adjusted values of the dependent variable for never-married dropouts, minus adjusted value for the relevant comparison group.

3. High school graduates may have some college, but have no Bachelor's degree.

4. Women in sample must have given birth to the same number of children as the number of own-children who live in the household.

5. See discussion beginning on page 2 for details concerning controls.

Table 3: Effect of state maximum welfare benefit on outmigration probability between 1975–80
(generating time: Fri Sep 19 19:24:18 2003 from file *outmigration-linear-marital-history-dropout.ara.*)

	Never-married dropouts			Ever-married dropouts			Never-married HS graduates			Ever-married HS graduates		
	[1]	[2]	[3]	[1]	[2]	[3]	[1]	[2]	[3]	[1]	[2]	[3]
Impact of max. ben.												
Oldest is 0 in 1975	-0.094	-0.265^{**}	-0.343^{***}	0.019	0.032	0.045	-0.104^{***}	-0.078	-0.053	-0.027	-0.052	-0.029
	(0.065)	(0.117)	(0.111)	(0.045)	(0.077)	(0.089)	(0.034)	(0.059)	(0.060)	(0.019)	(0.038)	(0.038)
× OC's age	0.002	0.039^{**}	0.044^{*}	-0.001	-0.006	-0.006	0.021^{***}	-0.003	-0.016	-0.001	-0.005	-0.007
	(0.008)	(0.018)	(0.023)	(0.006)	(0.011)	(0.013)	(0.008)	(0.015)	(0.012)	(0.002)	(0.005)	(0.005)
Marginal effects:												
Age 0–5	-0.009	-0.018^{*}	-0.024^{***}	0.003	0.003	0.005	-0.007^{*}	-0.011^{*}	-0.012^{*}	-0.006	-0.013^{*}	-0.009
	(0.006)	(0.009)	(0.008)	(0.006)	(0.010)	(0.012)	(0.004)	(0.006)	(0.006)	(0.004)	(0.007)	(0.008)
Age 6–11	-0.005	0.004	0.002	0.001	-0.003	-0.000	0.006	-0.010	-0.017^{**}	-0.006^{*}	-0.014^{*}	-0.013^{*}
	(0.004)	(0.007)	(0.008)	(0.003)	(0.007)	(0.008)	(0.005)	(0.009)	(0.007)	(0.004)	(0.007)	(0.008)
Baseline:												
Age 0–5	0.056			0.099			0.073			0.120		
Age 6–11	0.029			0.069			0.055			0.087		
N	2,345			4,008			3,531			16,097		
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Own-state controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
500-mile controls	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Note: * indicates significance at the 0.10 level; ** at the 0.05 level; *** at the 0.001 level. See text for other details.

Table 4: Effect of state maximum welfare benefit on outmigration probability between 1985–90
(generating time: Fri Sep 19 19:24:25 2003 from file *outmigration-linear-marital-history-dropout.ara.*)

	Never-married dropouts			Ever-married dropouts			Never-married HS graduates			Ever-married HS graduates		
	[1]	[2]	[3]	[1]	[2]	[3]	[1]	[2]	[3]	[1]	[2]	[3]
Impact of max. ben.												
Oldest is 0 in 1985	-0.051	-0.168***	-0.183***	0.011	-0.072	-0.098	-0.052*	-0.126***	-0.167***	-0.044*	-0.153***	-0.193***
	(0.033)	(0.048)	(0.066)	(0.038)	(0.085)	(0.088)	(0.031)	(0.048)	(0.055)	(0.025)	(0.057)	(0.055)
× OC's age	0.007	0.009	0.019*	-0.003	0.002	-0.002	0.007	0.001	0.009	0.000	0.007	0.008
	(0.007)	(0.008)	(0.011)	(0.006)	(0.011)	(0.011)	(0.005)	(0.006)	(0.009)	(0.002)	(0.005)	(0.005)
Marginal effects:												
Age 0–5	-0.004	-0.018***	-0.016***	0.001	-0.012	-0.018	-0.005	-0.017***	-0.020***	-0.008**	-0.025***	-0.032***
	(0.003)	(0.006)	(0.006)	(0.005)	(0.013)	(0.012)	(0.003)	(0.005)	(0.006)	(0.004)	(0.009)	(0.008)
Age 6–11	0.001	-0.009	-0.002	-0.002	-0.007	-0.015*	0.001	-0.012***	-0.010**	-0.006***	-0.013***	-0.018***
	(0.004)	(0.005)	(0.005)	(0.003)	(0.010)	(0.008)	(0.004)	(0.004)	(0.004)	(0.002)	(0.005)	(0.004)
Baseline:												
Age 0–5	0.063			0.103			0.074			0.105		
Age 6–11	0.043			0.071			0.053			0.077		
N	8,518			7,564			16,695			43,281		
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Own-state controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
500-mile controls	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Note: * indicates significance at the 0.10 level; ** at the 0.05 level; *** at the 0.001 level. See text for other details.

Table 5: Robustness checks on relationship between state maximum welfare benefit and outmigration probabilities
(generating time: Fri Sep 19 19:30:22 2003 from file *outmigration-linear-robustness.ara.*)

	Single college graduates			Married dropouts			Married HS graduates			Married college graduates		
	[1]	[2]	[3]	[1]	[2]	[3]	[1]	[2]	[3]	[1]	[2]	[3]
A. 1980 sample coefs:												
Oldest is 0 in 1975	0.104	0.205*	0.218*	-0.013	0.043	0.021	-0.030	-0.044	-0.058	0.010	0.049	0.048
	(0.085)	(0.110)	(0.114)	(0.027)	(0.048)	(0.060)	(0.025)	(0.046)	(0.047)	(0.046)	(0.052)	(0.055)
× OC's age	-0.007	-0.011	-0.011	0.004	-0.007	-0.004	0.000	-0.001	0.000	-0.002	-0.007**	-0.006
	(0.006)	(0.008)	(0.009)	(0.003)	(0.004)	(0.004)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
Baseline	0.1			0.1			0.1			0.2		
N	3,161			35,815			144,210			31,995		
B. 1990 sample coefs:												
Oldest is 0 in 1985	-0.022	-0.135***	-0.100	-0.052*	-0.089	-0.046	-0.061***	-0.107**	-0.154***	-0.059***	-0.082*	-0.052
	(0.027)	(0.040)	(0.069)	(0.030)	(0.067)	(0.056)	(0.019)	(0.042)	(0.041)	(0.021)	(0.043)	(0.042)
× OC's age	-0.005	0.005	0.007	0.001	0.003	-0.000	0.001	-0.000	0.001	-0.002*	-0.002	-0.004
	(0.003)	(0.004)	(0.006)	(0.001)	(0.003)	(0.003)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.003)
Baseline	0.1			0.1			0.1			0.1		
N	11,510			53,305			304,275			106,492		

Note: * indicates significance at the 0.10 level; ** at the 0.05 level; *** at the 0.001 level. See text for other details.

Table 6: Simulated effect of a 10% cut in maximum benefit(generating time: Sun Jun 29 14:54:33 2003 from file *simulation-results.ara.*)

	HS Dropouts			HS Graduates			Overall		
	$\tilde{r}_{.05}$	Point	$\tilde{r}_{.95}$	$\tilde{r}_{.05}$	Point	$\tilde{r}_{.95}$	$\tilde{r}_{.05}$	Point	$\tilde{r}_{.95}$
Mississippi									
$\tilde{r}_{H,lb}$	0.005	0.027	0.066	-0.002	0.001	0.096	0.002	0.021	0.045
$\tilde{r}_{H,ub}$	0.079	0.247	0.756	-0.063	0.716	19.709	0.029	0.365	0.712
Illinois									
$\tilde{r}_{H,lb}$	0.016	0.066	0.163	0.000	0.004	0.083	0.009	0.023	0.074
$\tilde{r}_{H,ub}$	0.038	0.151	0.348	0.001	0.071	0.345	0.033	0.095	0.228
Wisconsin									
$\tilde{r}_{H,lb}$	0.009	0.044	0.068	-0.008	0.031	0.097	0.005	0.035	0.069
$\tilde{r}_{H,ub}$	0.011	0.065	0.103	-0.015	0.049	0.142	0.004	0.054	0.104
N	987	987	987	941	941	941	931	931	931

Notes: Point estimates use actual data. Figures in columns headed $\tilde{r}_{.05}$ and $\tilde{r}_{.95}$ are the bootstrap replication that yielded the $.05^{th}$ and $.95^{th}$ percentiles based on 1000 bootstrap replications. See text for explanation of difference between $\tilde{r}_{H,lb}$ and $\tilde{r}_{H,ub}$.

Table 7: CLAD model estimates for never-married dropouts without and with migration variables
(generating time: Sat Jun 28 21:21:54 2003 from file *income-model-sensitivity.ara.*)

	Main model	Model with migration variables	Difference
Maximum benefit			
Age-0	1.51***	1.57***	0.06
	(0.37)	(0.59)	(0.64)
× oldest child's age	0.04	0.04	-0.00
	(0.08)	(0.30)	(0.31)
Five-year migration			
Age-0		-0.01	
		(42.15)	
× oldest child's age		0.00	
		(8.53)	
Change in maximum benefit			
Age-0		-0.43	
		(0.94)	
× oldest child's age		-0.05	
		(0.17)	
Mean of dependent variable	4.10		
Convergent replications	235	236	235

Note: Coefficient estimates are from models estimated using actual data. Figures in parentheses are standard errors calculated from 250 bootstrap replications (not all of which converged for both specifications). For difference column, standard errors are based on the bootstrap distribution of the difference in coefficients. See text for other right-hand side variables included in the models.

Table 8: Summary statistics for estimation samples
(generating time: Thu Jun 12 21:46:50 2003 from file *summary-statistics-revision.ara*.)

	1980 Sample				1990 Sample			
	Never married		Ever married		Never married		Ever married	
	Dropouts	HS Grads	Dropouts	HS Grads	Dropouts	HS Grads	Dropouts	HS Grads
White	0.23	0.24	0.72	0.80	0.29	0.34	0.72	0.83
	(0.42)	(0.43)	(0.45)	(0.40)	(0.45)	(0.47)	(0.45)	(0.38)
Hispanic	0.19	0.05	0.14	0.05	0.23	0.10	0.19	0.07
	(0.39)	(0.21)	(0.34)	(0.22)	(0.42)	(0.30)	(0.39)	(0.26)
Mother's age	28.5	29.1	31.5	32.5	29.0	30.4	31.9	33.8
	(5.4)	(4.8)	(6.2)	(5.0)	(5.0)	(5.0)	(5.6)	(5.1)
Youngest age, 80/90	5.0	6.3	7.1	8.0	5.1	6.4	7.0	8.2
	(3.9)	(3.8)	(4.1)	(3.7)	(4.0)	(4.2)	(4.2)	(3.9)
Youngest age, 75/85	2.4	2.9	3.5	3.8	2.6	3.2	3.6	4.1
	(2.5)	(2.6)	(2.9)	(2.9)	(2.7)	(3.0)	(3.1)	(3.1)
# Children, 80/90	2.5	1.8	2.4	1.9	2.5	1.8	2.3	1.8
	(1.2)	(0.9)	(1.2)	(0.9)	(1.2)	(1.0)	(1.1)	(0.8)
# Children, 75/85	1.8	1.4	2.0	1.7	1.8	1.4	1.9	1.6
	(1.0)	(0.7)	(1.1)	(0.8)	(1.0)	(0.7)	(0.9)	(0.8)
Age of oldest	9.2	8.9	11.0	10.7	9.6	9.4	11.1	10.9
	(3.3)	(3.0)	(3.4)	(3.3)	(3.3)	(3.3)	(3.4)	(3.3)
Moved	4.6	6.8	8.2	10.3	5.5	6.7	8.5	9.0
	(21.1)	(25.2)	(27.5)	(30.4)	(22.9)	(24.9)	(27.9)	(28.6)
Pub. assist. inc.	406	219	245	102	296	154	189	74
	(421)	(355)	(390)	(279)	(373)	(288)	(347)	(234)
Any pub. assist.	0.66	0.38	0.39	0.17	0.58	0.33	0.36	0.15
	(0.47)	(0.48)	(0.49)	(0.37)	(0.49)	(0.47)	(0.48)	(0.35)
Max benefit	561	562	545	563	494	493	480	493
	(132)	(127)	(131)	(127)	(121)	(112)	(116)	(110)
Medicaid expenditures	463	453	430	431	425	433	416	419
	(125)	(117)	(113)	(106)	(109)	(105)	(102)	(101)
Per capita income	16.8	16.7	16.3	16.6	19.5	19.5	18.9	19.2
	(2.2)	(2.1)	(2.1)	(2.0)	(3.1)	(3.0)	(2.8)	(2.7)
School expenditures	3.9	3.9	3.7	3.8	4.9	5.0	4.6	4.8
	(0.9)	(0.8)	(0.8)	(0.7)	(1.7)	(1.6)	(1.5)	(1.5)
Unemployment rate	7.46	7.52	7.37	7.34	6.33	6.25	6.32	6.21
	(1.25)	(1.34)	(1.38)	(1.41)	(0.95)	(1.00)	(0.94)	(0.97)
Minimum wage	6.10	6.10	6.09	6.09	4.77	4.75	4.75	4.75
	(0.07)	(0.07)	(0.03)	(0.03)	(0.20)	(0.18)	(0.13)	(0.14)
Marriageable men ratio	0.84	0.84	0.92	0.93	0.86	0.85	0.93	0.94
	(0.09)	(0.09)	(0.07)	(0.07)	(0.10)	(0.10)	(0.07)	(0.07)
N	2,345	3,531	4,008	16,097	8,525	16,737	7,578	43,408

Figure 1: Impact of a 10% Illinois benefit cut on probability of one-year migration and holdover probability, by oldest child's age