Supplemental Appendix

A Extensions of the Baseline Model: Discrete Labor Supply Choices, Heterogeneous Elasticity and Program Participation Cost

In subsection 3.1, the income/labor supply choice is assumed to be continuous, the income/labor supply elasticity e is held constant across agents, and perfect compliance (i.e. those eligible will participate in Medicaid/CHIP) is assumed. This subsection investigates the implications of relaxing these assumptions and shows that the qualitative predictions in the previous subsections still hold true.

Discrete Labor Supply Choices In the preceding subsection, agents' pre-tax income choice is assumed to be continuous, implying that agents are free to choose their hours and hence perfectly control their income. Obviously, this may not be a realistic restriction as per Ashenfelter (1980), Ham (1982), Kahn and Lang (1991), Altonji and Paxson (1992), Dickens and Lundberg (1993) and Chetty et al. (2011). In this subsection, I will first derive the theoretical prediction only allowing an agent finitely many hours choices.[1] The main implication is still that certain agents will lower their labor supply in order to claim benefit when a notch is introduced.

Because of the discrete labor supply restriction, Z in this section is written explicitly as wH where w is considered to be distributed smoothly among agents. For exposition purposes, I discuss the case when H can only vary along the extensive margin; that is, an agent can only work full time or not work at all. The general case where H is allowed to take on more than two values is analogous. Let $H=0$, $H=1$ denote the labor supply choice of not working and working full time respectively. If workers are constrained to only these two labor supply options, then the maximization problem becomes $\max_{H\in\{0,1\}} u(C,wH)$ subject to the budget constraint (2) where Z is replaced by wH, and we solve the maximization problem by considering the following two scenarios.

1. $w \leqslant \gamma$. An agent with potential monthly wage below the cutoff can claim benefits whether she works or not. In other words, the budget constraint she faces is only the segment to the left of γ: $C=(1-t)wH+g$. Consequently, maximizing utility involves the comparison of $u(g,0)$ and $u((1-t)w+g,1)$. To characterize the solutions, consider the agent of type $\bar{n}^l$ who is indifferent between choosing $H=0$ and $H=1$ at wage

[1] The working paper versions of Saez (2010), Saez (1999) and Saez (2002), address this extension in their simulation section but do not discuss the predictions from a theoretical perspective.

w.[2] Therefore, $\bar{n}^l$ solves

$$u(g,0) \equiv g = (1-t)w + g - \frac{n}{1+1/e}(\frac{w}{n})^{1+1/e} \equiv u((1-t)w+g,1)$$

which implies that $\bar{n}^l(w) = (\frac{w^{1+1/e}}{[(1-t)w](1+1/e)})^e$. Since $\frac{n}{1+1/e}(\frac{w}{n})^{1+1/e}$ is decreasing in n (i.e. the disutility of working is less for an agent with high n), agents with wage w and of type $n \geqslant \bar{n}^l(w)$ choose $H = 1$ and those with $n < \bar{n}^l(w)$ choose $H = 0$. Because of the quasilinear utility functional form, $\bar{n}^l(w)$ also characterizes the work choice in the absence of the transfer program.

2. $w > \gamma$. An agent with potential monthly wage above the cutoff is eligible for benefits only if she chooses not to work. The type of agent who is indifferent between working and not working at wage w equates $u(g,0)$ and $u((1-t)w,1)$. Because her type $\bar{n}^r$ solves

$$u(g,0) \equiv g = (1-t)w - \frac{n}{1+1/e}(\frac{w}{n})^{1+1/e} \equiv u((1-t)w,1)$$

$\bar{n}^r(w) = (\frac{w^{1+1/e}}{[(1-t)w-g](1+1/e)})^e$.[3] Analogous to case 1, agents with $n \geqslant \bar{n}^r(w)$ choose to work full time while those with $n < \bar{n}^r(w)$ choose not to work.

Let $\bar{n}_{0,1}$ denote the type of agents who are indifferent between working and not working, and it follows that $\bar{n}_{0,1}(w) = \begin{cases} \bar{n}^l(w) & \text{if } w \leqslant \gamma \\ \bar{n}^r(w) & \text{if } w > \gamma \end{cases}$. The threshold $\bar{n}_{0,1}$ varies smoothly with w for $w \leqslant \gamma$ and for $w > \gamma$, but there is a discontinuous increase in $\bar{n}_{0,1}$ as w crosses γ because $g > 0$. This means that certain workers will choose not to work when a notch is introduced if n and w have a smooth joint distribution $f_{n,w}$ supported over the first quadrant of $\mathbb{R}^2$.

Heterogeneous labor supply elasticities The qualitative prediction of the model holds true when elasticities are heterogeneous across families. In section 3.1, the threshold taste parameter, $\bar{n}$, is a function of e as per (3), and all statements are true for each $e > 0$. Denote this type threshold by $\bar{n}(e)$, let the conditional c.d.f. of n given e by $F_{n|e}$, and suppose that e is distributed smoothly across agents with a p.d.f. of f_e. The fraction of agents that lower their labor supply when a benefit notch is introduced is simply $\int_0^\infty (F_{n|e}(\bar{n}(e)|e) - F_{n|e}(n_\gamma(e)))f_e(e)de$ where $n_\gamma(e) \equiv \frac{\gamma}{(1-t)^e}$ as defined in section 3.1. This fraction is posi-

[2] The superscript l here stands for left as w lies to the left of γ. The superscript r will be used in the next case.

[3] Note that a positive n^r exists – n^r has to be positive for the marginal utility of work to be negative – when $(1-t)w > g$, which means that the post-tax income of working full time at wage w is larger than the value of benefit g. This is most likely satisfied for families with a wage above the CHIP cutoff.

tive because $\bar{n}(e) > n_\gamma(e)$ for $e \geqslant 0$.

Non-participation To account for non-participation among eligible agents, I follow a conventional approach by Moffitt (1983) and introduce a cost term to program participation. The cost term can encapsulate the simple psychological cost of being perceived as a beneficiary of government programs, but also the time and monetary cost of applying for benefits, such as filling out the required forms and acquiring program-related information. The simplest formulation is to add a flat cost to the utility function if the agent decides to participate in the program:

$$\max_{C,Z,P} u(C,Z) - \phi P$$

where an agent's welfare participation decision $P \in \{0,1\}$ depends on the cost parameter $\phi > 0$.

In effect, introducing cost shifts down the program segment of the budget constraint $[Z(1-t)+g]1_{[Z \leqslant \gamma]}$ by ϕ and therefore reduces the public insurance notch to $\max\{g-\phi, 0\}$. If ϕ is constant across agents and $\phi < g$, then all the analyses in subsection 3.1 carry through by replacing g with $\tilde{g} = g - \phi$. When ϕ is heterogeneous, the income distribution is smooth for the sub-population with $\tilde{g} = g - \phi \leqslant 0$, and analyses from previous subsections only hold true for those with $\tilde{g} > 0$. In the entire population, the qualitative predictions from subsection 3.1 are still valid if (n, e, ϕ) follows a smooth distribution supported on $\mathbb{R}^4_{++}$.[4]

B Dynamic Labor Supply Model in Subsection 3.2

In this section, I provide details for the dynamic labor supply model in subsection 3.2. Formally, the state variable s is the number of months until recertification (s is defined to be 0 for those not claiming benefits since they will face the eligibility check when they apply). Let τ be the number of months of provided continuous eligibility. In each period, an agent chooses whether to participate in the program:

$$V_s = \max_{P_s} P_s V_s^1 + (1-P_s) V_s^0 \tag{9}$$

where $P_s = 0, 1$ denotes participation choice, and V_s^1 and V_s^0 are utilities associated with participating and not participating in the program when agents are s months away from an eligibility check. The expressions for V_s^1 and V_s^0 are

[4]Note the same argument applies if heterogeneity in g is allowed, and the model's prediction still holds true in that case. Heterogeneity in g may be expected because families with healthier children value health insurance less than those with sicker children, for example.

$$V_s^1 = \max_{C,Z}\{u(C,Z)+\beta V_{s'}\}$$
$$\text{s.t. } Z<\gamma \text{ if } s=0;\ C=(1-t)Z+g$$
$$s' = \begin{cases} s-1 & \text{if } s>0 \\ \tau-1 & \text{if } s=0 \end{cases}$$

$$V_s^0 = \max_{c,z}\{u(C,Z)+\beta V_{s'}\}$$
$$\text{s.t. } C=(1-t)Z$$
$$s' = \begin{cases} s-1 & \text{if } s>0 \\ 0 & \text{if } s=0 \end{cases}$$

I introduce the notation $\{C_s^p, Z_s^p\} = \text{argmax} V_s^p$ for $p=0,1$, and the dynamic problem simplifies to

$$\begin{aligned} V_0 &= \max_{P_0} P_0\{u(C_0^1,Z_0^1)+\beta V_1\}+(1-P_0)\{u(C_0^0,Z_0^0)+\beta V_0\} \\ V_1 &= \max_{P_1} P_1\{u(C_1^1,Z_1^1)+\beta V_0\}+(1-P_1)\{u(C_1^0,Z_1^0)+\beta V_0\} \end{aligned} \tag{10}$$

I characterize the optimal P_s, C_s^p and Z_s^p's below for $s=0,1$ and $p=0,1$.

First note that choosing $P_1=1$ strictly dominates $P_1=0$ because (C_1^0,Z_1^0) lies in the interior of the budget set for an agent with $s=1$. In other words, when benefits can be claimed without having to lower income, a rational family will do so. This reasoning simplifies the expression for V_1 to $V_1 = u(C_1^1,Z_1^1)+\beta V_0$. Plugging in this expression of V_1 into that of V_0 leads to

$$V_0 = \max_{P_0} P_0\{u(C_0^1,Z_0^1)+\beta u(C_1^1,Z_1^1)+\beta^2 V_0\}+(1-P_0)\{u(C_0^0,Z_0^0)+\beta V_0\}$$

For the agents indifferent between choosing $P_0=0$ and $P_0=1$,

$$V_0 = u(C_0^1,Z_0^1)+\beta u(C_1^1,Z_1^1)+\beta^2 V_0 = u(C_0^0,Z_0^0)+\beta V_0$$

and therefore $V_0 = \frac{u(C_0^0,Z_0^0)}{1-\beta}$. It follows that

$$u(C_0^1,Z_0^1)+\beta u(C_1^1,Z_1^1) = u(C_0^0,Z_0^0)+\beta u(C_0^0,Z_0^0) \tag{11}$$

For quasi-linear utility u, $C_1^1 = C_0^0+g$ and $Z_1^1 = Z_0^0$. Consequently, $u(C_1^1,Z_1^1) = u(C_0^0,Z_0^0)+g$, and (11) leads to

$$u(C_0^1,Z_0^1)+\beta g = u(C_0^0,Z_0^0) \tag{12}$$

Suppose (C_0^1, Z_0^1) satisfying (12) is an interior solution. Then the convex indifference curve passing through the bundle (C_0^1, Z_0^1) is tangent to the program segment of the budget constraint and therefore lies above the non-program budget constraint $C = (1-t)Z$. Consequently, $u(C_0^1, Z_0^1) > u(C_0^0, Z_0^0)$ implying that $u(C_0^1, Z_0^1) + \beta g > u(C_0^0, Z_0^0)$, contradicting (12). Therefore, the (C_0^1, Z_0^1) that satisfies (12) has to be a corner solution with $Z_0^1 = \gamma$. Denote the indifferent agent's type by $\bar{n}^{dynamic}$ and expanding (12) using the quasi-linear functional form leads to

$$\gamma(1-t) + (1+\beta)g - \frac{\bar{n}^{dynamic}}{1+1/e}(\frac{\gamma}{\bar{n}^{dynamic}})^{1+1/e} = \bar{n}^{dynamic}(1-t)^{1+e} - \frac{\bar{n}^{dynamic}}{1+1/e}(1-t)^{1+e} \quad (13)$$

Equation (13) states that an agent of type $\bar{n}^{dynamic}$ is indifferent between choosing her interior solution on the budget constraint segment $C = (1-t)Z1_{\{Z>\gamma\}}$ and the post-tax/pre-tax income bundle ($\gamma(1-t) + (1+\beta)g$, γ).

C Seam-bias Adjustment in the Calibration Exercises

In this section, I describe the seam-bias adjustment used in section 6.1. First, to see why seam-bias may dampen the observed strategic behavior, let T^* and T denote the true and reported reference month of the beginning of a public insurance spell; since there are four reference months within a wave, T^* and T both take on the values of 1 through 4. Let Y_m^* and Y_m denote the true and reported family income for month m of the public insurance spell.[5] As an example, consider a child who started her spell in October and whose family is interviewed by SIPP in December. Since September, October, November and December will be reference months 1 through 4 for her family, her true transition reference month of October implies that $T^* = 2$. If her family telescopes, public insurance coverage will be reported for the entire wave based on the fact that she is covered in December, and the reported reference month for starting public insurance is September, implying $T = 1$. Telescoping in income suggests that her family's reported September income is the same as their December income. Because September is month 1 in the reported public insurance spell and December is month 3 in the true public insurance spell, it follows that $Y_1 = Y_3^*$. By the same logic, $Y_0 = Y_{-1}^*$, and the observed rebound magnitude reported for spell month 1 is $\Delta Y_1 \equiv Y_1 - Y_0 = Y_3^* - Y_{-1}^*$,

[5]For exposition, I only focus on the measurement error induced by the seam bias here and assume that families report their true income in the absence of telescoping behavior. Adding a standard classical measurement error component will not affect the analysis since it will be averaged out when examining the mean income responses.

which is zero if the family behaves according to a model with no income effect. Applying the same line of reasoning reveals that ΔY_1 is a valid measure of the model-predicted rebound magnitude only for those with $T^* = 1$ among the telescoping families.

To address the complications resulting from the seam bias, I adopt some of the assumptions proposed in Ham et al. (2009):[6]

Assumption 1) A respondent either telescopes or reports truthfully: $T = 1$ or $T = T^*$,

Assumption 2) The true reference month for starting a public insurance spell has a uniform distribution, implying $\Pr(T^* = t^*) = \frac{1}{4}$ for $t^* = 1,2,3,4$.

It follows from Assumption 1) that if a transition into public insurance truly happened in reference month 1, then the respondent reports so: $T^* = 1 \Rightarrow T = 1$. It also implies that if a child's public insurance spell was reported to start in the second, third or fourth reference month, then it truly started in that month: $T = s \Rightarrow T^* = s$. Assumption 2) is justified by the SIPP survey design as noted in Ham et al. (2009). The entire sample is randomly split into four rotation groups, and one rotation group is interviewed each calendar month. Therefore, each calendar month is reference month 4 for rotation group 1, reference month 3 for rotation group 2, reference month 2 for rotation group 3 and reference month 1 for rotation group 4.

If all families behave according to the labor supply model with parameter ρ and e and the corresponding rebound magnitude $y^{\rho,e}$, then for $k = 1,...,11$,

$$\begin{aligned} E[\Delta Y_k] &= \sum_{s=1}^{4} E[\Delta Y_k | T = s] \Pr(T = s) \\ &= \underbrace{E[\Delta Y_k | T = 1] \Pr(T = 1)}_{(1)} + \underbrace{\sum_{s=2}^{4} E[\Delta Y_k | T = s] \Pr(T = s)}_{(2)} \end{aligned}$$

[6] Assumptions 1) combines A1), A2) and A4) , and Assumption 2) is in the spirit of A5) of Ham et al. (2009). Unlike Ham et al. (2009), however, I do not attempt to impose parametric restrictions and estimate the probabilities of $\Pr(T = 1 | T^* = t^*)$ for $t^* = 2,3,4$.

Examining the two terms separately:

$$
\begin{aligned}
(1) &= \{\sum_{s^*=1}^{4} E[\Delta Y_k|T=1,T^*=s^*]\Pr(T^*=s^*|T=1)\}\Pr(T=1) \\
&= E[\Delta Y_k|T=1,T^*=1]\Pr(T^*=1|T=1)\Pr(T=1) \\
&= E[\Delta Y|T=1,T^*=1]\Pr(T=1|T^*=1)\Pr(T^*=1) \\
&= E[\Delta Y|T=1,T^*=1]\Pr(T^*=1) \\
&= \frac{1}{4}y^{\rho,e}
\end{aligned}
$$

where the second equality is a result of the fact that the expected rebound magnitude is 0 for $s^*=2,3,4$ and the fourth and fifth equalities follow from Assumptions 1) and 2) respectively. By Assumption 1)

$$
\begin{aligned}
& E[\Delta Y_k|T=s]\Pr(T=s) \\
=& \sum_{s^*=1}^{4} E[\Delta Y_k|T^*=s^*,T=s]\Pr(T^*=s^*|T=s)\Pr(T=s) \\
=& E[\Delta Y_k|T^*=s^*,T=s]\Pr(T=s) \\
=& y^{\rho,e}\Pr(T=s)
\end{aligned}
$$

for $s=2,3,4$. It follows that

$$\frac{E[\Delta Y_k]}{\frac{1}{4}+\sum_{s=2}^{4}\Pr(T=s)} = y^{\rho,e}$$

Under these assumptions and the null hypothesis that the families behave according to the labor supply model, the following relationship is established linking the observed average rebound in spell month k ($k=1,...,11$), $E[\Delta Y_k]$, to the predicted magnitude $y^{\rho,e}$ from a model with curvature parameter ρ and elasticity e:

$$\frac{E[\Delta Y_k]}{\frac{1}{4}+\sum_{s=2}^{4}\Pr(T=s)} = y^{\rho,e} \tag{14}$$

The intuition of equation (14) becomes clear when it is applied in two special cases. If no one telescopes, $\sum_{s=2}^{4}\Pr(T=s)=\frac{3}{4}$ and consequently $E[\Delta Y_k]=y^{\rho,e}$. If everyone telescopes, on the other hand, seam bias is at its worst as all spells are reported to begin at the beginning of a wave and consequently $E[\Delta Y_k]=\frac{1}{4}y^{\rho,e}$, confirming the aforementioned attenuation bias.

In practice, equation (14) may not hold for $k=9,10,11$, because the timing of the strategic behavior for families facing eligibility recertification is ambiguous as discussed in section 2 and that telescoping affects these three spell months differently even if the recertification process requires income proof from month 12.[7] Therefore, the statistical tests in section 6 only rely on equation (14) for $k = 1, ..., 8$.

Denoting the left hand side of (14) by $E[\Delta Y_k^{adj}]$, the equation says that the observed rebound adjusted for seam bias $E[\Delta Y_k^{adj}]$ should be equal to the calibrated magnitudes if families behave according to the labor supply model. Since (14) holds for all $k = 1, ..., 8$, it also holds for the average of $E[\Delta Y_k^{adj}]$ across k (denoted by y^{emp}). The statistical analysis in section 6 involves testing the null hypothesis of H_0: $y^{emp} = y^{\rho,e}$ against the one-sided alternative H_1: $y^{emp} < y^{\rho,e}$.

To implement the test, I estimate the empirical rebound y^{emp} and calculate the model-calibrated rebound $y^{\rho,e}$ in each bootstrap sample. In particular, I use $\frac{1}{N}\sum_i 1_{[T_i=s]}$ (N is the number of public insurance spells as well as the number of spells in each bootstrap sample) as the sample analog for $\Pr(T = s)$ and $\hat{\delta}_k$ from the regression of (5) as the estimator for $E[\Delta Y_k]$. The implied estimator for y^{emp} is $(\frac{1}{8}\sum_{k=1}^{8}\hat{\delta}_k)/(\frac{1}{4} + \frac{1}{N}\sum_i\sum_{s=1}^{3} 1_{[T_i=s]})$ in each bootstrap sample, and the estimator for $y^{\rho,e}$ is simply the average of the calibrated rebound magnitudes across public insurance spells.

D Tests Using Counterfactual Groups: Details

In this section, I describe the symmetric difference-in-differences estimation used in the first counterfactual analysis in section 6.2. Following Ashenfelter and Card (1985), if the shocks in the income process are joint normal, then the conditional expectation of income in spell month k is a linear function of that in month r:

$$E[Y_{ik}|Y_{ir} < \tilde{y}] = \omega_i + \sum_t D_{it}^k \lambda_t + \xi_k + \frac{Cov(Y_{ik}, Y_{ir})}{Var(Y_r)}\{E[Y_{ir}|Y_{ir} < \tilde{y}] - E[Y_{ir}]\} \quad (15)$$

$$E[Y_{jk}|Y_{jr} \geqslant \tilde{y}] = \omega_i + \sum_t D_{jt}^k \lambda_t + \xi_k + y_k E[O_i] + \frac{Cov(Y_{jk}, Y_{jr})}{Var(Y_r)}\{E[Y_{jr}|Y_{jr} \geqslant \tilde{y}] - E[Y_{jr}]\} \quad (16)$$

Because of the existence of the last term in (15) and (16), which reflects serial correlation in transitory shocks, running fixed-effect regressions separately for the high and low income groups and taking the differences in the estimated $\hat{\delta}_k$'s will not in general consistently estimate the "intent-to-treat" effect $y_k E[O_i]$. Assuming covariance stationarity in the income process, however, the symmetric difference-in-differences

[7] In this case, (14) changes to $\frac{E[\Delta Y_k]}{\frac{1}{4}\Pr(T^*=1|T=1)+\sum_{s=2}^{4}\Pr(T=s)} = y^{\rho,e}$, which depends on the unknown fraction of the nontelescopers among those with $T = 1$.

centered around month r for children who started their public insurance spell in the same calendar month, $SDD_r \equiv (E[Y_{jk}|Y_{jr} \geqslant \tilde{y}] - E[Y_{j0}|Y_{jr} \geqslant \tilde{y}]) - (E[Y_{ik}|Y_{ir} < \tilde{y}] - E[Y_{i0}|Y_{ir} < \tilde{y}])$ where $0 < 2r = k < 12$, identifies the "intent-to-treat" income rebound $(y_k - y_0)E[O_i]$.

In practice, I carry out the estimation with $r = 3$ and $r = 4$.[8] As mentioned in section 6.1, the maximum k for gauging income rebound needs to be capped at $k = 8$, implying a maximum r of 4. The values of $r = 1$ and $r = 2$ are not selected because the presence of seam bias threatens covariance stationarity.[9] To implement the test, I estimate each of the differences $E[Y_{jk}|Y_{jr} \geqslant \tilde{y}] - E[Y_{j0}|Y_{jr} \geqslant \tilde{y}]$ and $E[Y_{ik}|Y_{ir} < \tilde{y}] - E[Y_{i0}|Y_{ir} < \tilde{y}]$ by fixed effect regression, accounting for the seam bias using spells that began in the same calendar month t, and the estimator for SDD_r is averaged across t. Testing the labor supply model amounts to testing H_0: $SDD_r = 0$ versus $H_1 : SDD_r > 0$, and the number of bootstrap iterations is determined according to Andrews and Buchinsky (2000).

E Optimal Length of the Continuous Eligibility Period: Details

E.1 Further discussion of the social welfare function in Section 7

The formulation of the social welfare function (6) differs from a textbook approach (e.g., Salanie (2003)) in the following two respects. First, government surplus does not typically enter a social welfare function directly but through a balanced budget constraint. As noted by Salanie (2003), however, the dependence of utility on S is omitted in a textbook model because the spending on the public good is held constant. The specification (6) simply extends that of Salanie (2003) by allowing the production of the public good to be variable.

Second, having a "notched" lump sum transfer schedule with the associated cutoff γ as the policy instrument is not prevalent in the optimal design literature. In fact, if the income tax schedule is completely flexible and that $\Psi \circ u$ is strictly concave, then the government should choose a transfer function that equalizes consumption across agents when labor supply decisions are not considered in the model (a special case is studied as early as in Edgeworth (1897)). When labor supply incentives are considered, the seminal paper Mirrlees (1971) shows that the marginal tax rate always lies between zero and one, which precludes a discrete drop in the consumption-pre-tax-income schedule if the optimal tax schedule is completely flexible.

[8] Since the results are similar, I only report that for $r = 4$ in Table 7.

[9] Recall that telescoping families start their spell at the beginning of a wave, and therefore the covariance $Cov(Y_{i0}, Y_{ir})$ is cross-wave whereas $Cov(Y_{ik}, Y_{ir})$ is within-wave for $r = 1, 2$.

However, Blinder and Rosen (1985) and Slemrod (2010) argue that it is possible to institute a notch as part of an optimal schedule when the set of income tax instruments is limited, e.g., linear.[10] By continuing with the specification of (6), I take as given the existence of the notch-creating transfer programs like Medicaid and CHIP.

E.2 Optimal Recertification Frequency: Additional Results

This subsection presents additional results from the optimal continuous eligibility period calculations. First, I investigate the optimal recertification frequency under partial take-up. The point-in-time participation rate among eligibility children is estimated to be 82% in 2008 by Kenney et al. (2011), which is the value I use in my calculations below.

Table SA.1 presents the optimal length of the continuous eligibility period from simulations under 25 combinations of κ and ϕ – five values of \$0, \$9.5, \$19, \$28.5 and \$38 for both κ and ϕ, three values of η, $\eta = 0, 0.5, 1$[11] and two assumptions governing the take-up rates. The prevalence of the optimal continuous eligibility periods that are multiples of 4 under 100% take-up is again due to the seam bias; it is not so for the partial take-up results because of the random monthly program participation introduced. Under 100% take-up (column blocks (a) and (c)), the optimal recertification frequency is indeterminate for the utilitarian government ($\eta = 0$) when monitoring is costless because transferring wealth across population leads to no change in the overall welfare. When monitoring is costly, any eligibility check imposes a deadweight loss and therefore the implied optimal interval is the corner solution of 35 months. For the concave social welfare functions considered ($\eta > 0$), the optimal τ is smaller because of the pressure to efficiently target the needy. As with the patterns in 8, an increase in the cost on families, ϕ, is more likely to lengthen the recertification period than an increase in κ of the same magnitude. Under partial take-up (columns block (b) and (d) of Table 8), the calculated optimal τ's are no longer monotone in the cost parameters because of the randomness in take-up behavior in the simulations.[12] As mentioned in section 7, the recertification periods

[10] The *theoretical* properties of means-tested in-kind transfers in an optimal-design context have also been studied in Nichols and Zeckhauser (1982), Blackorby and Donaldson (1988), Gahvari (1995), Cremer and Gahvari (1997), Singh and Thomas (2000), etc. These studies typically consider the problem with two types of agents and a transfer scheme that ensures second-best allocation, i.e. the high type does not pretend to be the low type and claims benefit transfer. See Currie and Gahvari (2008) for a survey.

[11] $\eta = 0$ implies a linear (utilitarian) Ψ. Under the four values of η, giving \$100 per month to a family with a monthly income of \$1000 brings the same increment to social welfare as giving \$100, \$311 and \$1000 to a family with a monthly income of \$10,000, respectively. Larger values of η are also tried, which imply longer recertification periods.

[12] Also due to randomness, the optimal τ when $\eta = 0$ and $\phi, \kappa > 0$ is not necessarily the corner solution of 35 months. It could happen by chance in the simulation that no family participates in public insurance when $\tau = 34$ but all eligible families participate when $\tau = 35$, in which case the social welfare from $\tau = 35$ is lower because of the deadweight loss of monitoring.

are longer under partial take-up because short continuous eligibility periods lead to coverage gaps, reducing welfare for low-income families.

Next, I investigate the sensitivity of the normative results to alternative sample restrictions. Specifically, I carry out the same exercise in section 7 but use families with children who did not appear for the entire panel. The two alternative samples are: a) the "24-month sample": families with children who appeared consecutively for at least 24 months (i.e., the number of consecutive appearances between 24 and 35 months for the 2001 panel, and between 24 and 47 months for the 2004 panel); b) the "12-month sample": families with children who appeared consecutively for at least 12 months but no more than 23 months. To contrast with these alternative samples, I will refer to the sample used in section 7 as the "full-panel sample".[13]

Using the 24-month sample leads to very similar results as those in Table 8, and the optimal continuous eligibility period is at least 12 when $\phi \geqslant \$19$. In comparison, the 12-month sample yields somewhat shorter eligibility periods: when $\phi \geqslant \$19$, the optimal recertification period is eight months assuming full take-up. The reason for the shorter optimal period in the 12-month sample is that there appear to be more eligibility threshold crossing in that sample: the average number of times over a 12-month period that family income crosses the Medicaid eligibility threshold is 1.90 and 1.53 for the 12-month samples in the two SIPP panels; the corresponding numbers are 0.90 and 0.75 for the full-panel samples. The crossings in the 12-month sample call for more frequent recertification to improve targeting efficiency.

However, assuming partial take-up pushes the continuous eligibility period to the corner solution of 11 months in the 12-month sample. In addition, the 12-month sample only accounts for about one third of the families in the three samples. Overall, the 12-month lower bound for the continuous eligibility period is not very sensitive to alternative sample restrictions.

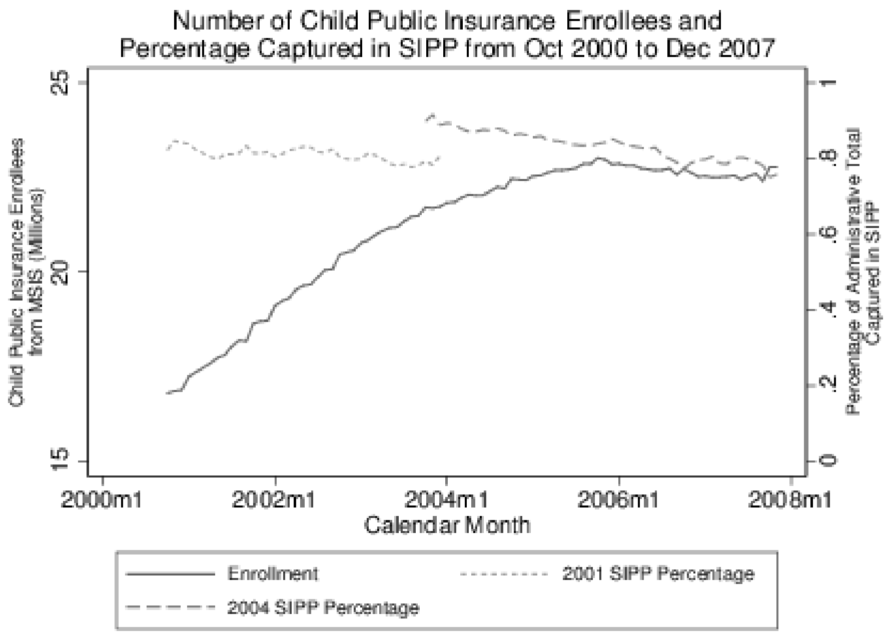
E.3 Comparison to Prell (2008)

The normative framework presented in this study relates to and extends the informative Prell (2008) model in studying the optimal WIC recertification frequency along several major dimensions. First, Prell (2008) assumes constant hazard rates in the transitions between eligibility and ineligibility, which makes the problem analytically tractable and provides nice insights. In comparison, I carry out the exercise non-parametrically by relying on the empirical distribution of incomes and provide a computational solution. This approach also allows the incomes to endogenously respond to tax rates, transfer notches as long as they do not re-

[13] Together, the three samples capture over 70 percent of families who had an infant at some point during the panel.

spond to the eligibility recertification period. Second, Prell (2008) assumes 100% program participation but acknowledges that take-up behavior or "program access" should be modeled. Estimating the take-up probability and building it into the normative framework brings this analysis a step closer to the goal. Third, the value of the transfer to different individuals is assumed to be the same from the social planner's perspective in the Prell (2008) framework whereas I calculate the optimal recertification interval under alternative social welfare functions. Using my framework, the implied optimal continuous eligibility periods under the 100% take-up rate are moderately longer than those of Prell (2008) in the WIC context, while those under partial take-up are much larger.

Figure SA.1: Number of Child Public Insurance Enrollees and Percentage Captured in SIPP



Note: The solid line with the left y-axis represents the total number public insurance enrollees per month between October 2000 and December 2007 who were eligible as dependent children; the underlying data are extracted from the Medicaid Statistical Information System. The two dashed lines with the right y-axis represent the number of child enrollees estimated from the 2001 and 2004 panels of SIPP, respectively, as a percentage of the administrative total (the solid line). The deviation of the dashed lines from the value of 1 reflect the degree of under-reporting.

Table SA.1: Optimal Length of the Continuous Eligibility Period: Other Specifications

Recertification Cost		Optimal Length of the Continuous Eligibility Period in Months (τ)											
ϕ	κ	SIPP 2001 Panel						SIPP 2004 Panel					
		(a) Full Take-up			(b) Partial Take-up			(c) Full Take-up			(d) Partial Take-up		
		$\eta = 0$	0.5	1	$\eta = 0$	0.5	1	$\eta = 0$	0.5	1	$\eta = 0$	0.5	1
$0	$0	1-35	1	1	1-35	6	22	1-35	1	1	1-35	8	25
$0	$9.5	35	4	4	35	9	20	35	4	4	35	12	24
$0	$19	35	4	4	35	9	21	35	4	4	35	8	32
$0	$28.5	35	4	4	35	12	20	35	8	4	35	13	17
$0	$38	35	8	4	35	12	21	35	8	4	35	12	18
$9.5	$0	35	4	4	35	9	19	35	4	4	35	8	25
$9.5	$9.5	35	4	4	35	12	19	35	4	4	35	12	25
$9.5	$19	35	4	4	35	12	18	35	8	8	35	12	24
$9.5	$28.5	35	8	8	35	12	18	35	8	8	35	16	26
$9.5	$38	35	8	8	35	12	18	35	12	8	35	12	25
$19	$0	35	4	8	35	12	18	35	8	8	35	12	26
$19	$9.5	35	8	8	35	12	22	35	8	8	35	12	25
$19	$19	35	8	8	35	12	24	35	12	12	35	12	31
$19	$28.5	35	12	8	35	12	19	35	12	12	35	16	26
$19	$38	35	12	12	35	12	23	35	12	12	35	12	24
$28.5	$0	35	8	12	35	12	18	35	8	12	35	12	26
$28.5	$9.5	35	8	12	35	12	21	35	12	12	35	12	24
$28.5	$19	35	12	12	35	12	20	35	12	12	35	12	24
$28.5	$28.5	35	12	12	35	13	19	35	12	12	35	17	25
$28.5	$38	35	12	12	35	19	20	35	12	12	35	16	24
$38	$0	35	12	12	35	12	19	35	12	12	35	16	26
$38	$9.5	35	12	12	35	12	19	35	12	12	35	12	28
$38	$19	35	12	12	35	18	18	35	12	12	35	12	25
$38	$28.5	35	12	12	35	18	24	35	12	12	35	17	35
$38	$38	35	12	12	35	20	20	35	16	16	35	25	29

Note: The optimal lengths of the continuous eligibility period, τ, are calculated based on the framework in section 7. The choice set of τ is $\{1, 2, ..., 35\}$. The calculation is carried out for different recertification costs (in 2010 dollars), different values of the social welfare function parameter η and different assumptions governing take-up rates. The SIPP sample that serves the basis of the calculation consists of families with children who had appeared every month during the 2001 and 2004 panels.