



La Follette Policy Report

Robert M. La Follette School of Public Affairs

Volume 16, Number 2, Fall 2006

Director's Perspective

La Follette welcomes APPAM, launches initiatives

This fall marks a special time for the Robert M. La Follette School of Public Affairs at the University of Wisconsin-Madison. We have begun several initiatives, including a weekly lunch seminar. If you are close to campus, we'd love to have you drop in. Or, if you would like to be on the e-mail list announcing each talk, just let us know via info@lafollette.wisc.edu.

We have launched a number of externally supported research projects this fall. One concerns major issues in higher education, especially the drift of many flagship state universities toward less economically diverse student bodies. This trend—described by some as the growing elitism of state-supported higher education—has been the subject of recent articles in the *New York Times*, *The Wall Street*

continued on page 14

In This Issue

Childhood Interventions that May Lead to Increased Economic Growth.....	1
Welfare and Child Support Policy: Who Should Benefit When Child Support is Paid to TANF Families.....	3
The Debate about the Undervaluation of China's Renminbi.....	15

Childhood Interventions that May Lead to Increased Economic Growth

Barbara Wolfe and Nathan Tefft

Barbara Wolfe is the La Follette School's new director and a professor of public affairs, economics, and population health sciences, at the University of Wisconsin-Madison. Supported by a grant from the National Institute of Mental Health, Nathan Tefft is a trainee with the Department of Economics at the University of Wisconsin-Madison. The research for this article was funded by The Pew Charitable Trusts, in collaboration with Robert Dugger and the Invest in Kids Working Group, and by a grant from the National Institutes of Health.

Childhood interventions have the potential to significantly affect economic growth in positive ways. Because childhood interventions, especially those for children 5 and younger, condition all future gains in human and health capital, they are likely to yield a high return in terms of economic benefits. The benefits of childhood interventions carry over to society as a whole through pathways such as increases in the skill of the labor force, which can attract employers and increase innovation; improved health of the population, which increases productivity and decreases sick time, thus improving work output while lessening demand on the health-care system; and reductions in crime, which reduces the burden on the criminal justice system.

Indicators of economic growth include increases in the level of schooling of young adults; earnings and total compensation and their distribution within the population; probability of labor force participation and employment; the number of healthy and productive work days; the level of productivity per hour of work; consumer efficiency in the purchase and use of consumer goods and services; and human capital of the next generation.

The impact of a program on these or any related indicators can be translated into changes in the rate of economic growth, a relationship many researchers have examined, especially in the context of early childhood interventions. From an economic and policy perspective, analysis of interventions ranging from prenatal care to programs for children in elementary school makes sense, since childhood is an important developmental stage in a person's life. Many follow-up studies of early childhood interventions show lasting effects in terms of greater academic achievement, higher

cognitive test scores, improved behavior patterns, lower crime initiation and recidivism, and even increased employment and wages, all indicators of economic growth. Not only do the effects last, interventions for younger children often produce stronger results than those implemented later in a child's development. These findings are consistent with economic models of human and health capital development, and they serve as promising indications of the possibility for enhancing the success and quality of life for many children, which should help improve the nation's economy. This evidence is more sentimentally captured in the phrase "our children are our future." Prosaically, if individuals' wages increase, their consumption and/or capital investment are likely to increase, thus fueling economic growth. At the same time, a more educated workforce is likely to increase productivity and the rate of technological innovation, which also leads to positive effects on economic growth.

For a policy-making body or a private organization, the challenge is to decide which childhood interventions are the best investments for economic growth. We offer criteria to help researchers identify programs that show promise. We use data from about 200 studies of child interventions that could be analyzed in terms of their contribution to economic growth. Our main emphasis is on experimental programs, since they tend to have stronger research promise, but we also include other programs. We do not restrict our attention to interventions for children of a certain age. Since there is evidence that interventions targeted at younger children appear to be particularly promising, we focus much of our attention on those programs. But we also consider programs targeted toward older children and adolescents. Our topic areas are arranged in roughly an age order with those targeted at the youngest children first and those for late adolescents last. Features of the programs under consideration include home visits by nurses; education, health care, and family services for preschool-age children; tutoring for elementary school children; health insurance for low-income families; teenagers being matched with adult mentors; efforts to keep students in high school and prepare them for what follows; and job training and other services. Table 1 details seven programs, five of which we discuss in more detail.

A successful childhood intervention implementation and evaluation must be designed so that causal mechanisms are identifiable. Researchers need to be able to say with confidence that confounding effects such as omitted variables or selection effects do not bias their results. In our review, we look for childhood

interventions that most closely achieve this identifiability. Randomization into a treatment group across as many characteristics as possible, or as is appropriate, is of course the most important criterion in this pursuit. To the extent that an intervention uses a truly random sample, we can be confident that no selection issues exist insofar as comparisons between treatment and control groups are concerned.

Next, the control groups must be constructed in such a way that the relevant characteristics of the childhood intervention are accurately separated out. In some cases, there may be effects linked to the treatment group that are secondary to the intent of the intervention but that actually drive the results. For example, a well-constructed control group might consist of children who are admitted to an intervention

continued on page 8

**The Robert M. La Follette
School of Public Affairs**

University of Wisconsin-Madison
1225 Observatory Drive
Madison, WI 53706
(608) 262-3581
Fax: (608) 265-3233
policyreport@lafollette.wisc.edu
www.lafollette.wisc.edu

Staff

Barbara Wolfe Director
Karen Holden Associate Director
Craig Allan Information Systems Manager
Karen FASTER Publications Director
Andrew Lambert Information Technology Specialist
Gregory Lynch Faculty Assistant
Elisabeth McInnis Career Development Coordinator
Mary Mead Student Services
Bridget Pirsch Outreach Specialist
Barbara Prigge Financial Specialist
Andrew Reschovsky *Policy Report* Editor
Kari Reynolds Administrator
Terry Shelton Outreach Director
Mary Treleven Student Services Coordinator

Centers of the La Follette School

Center for Public Policy and the Status of Women
Center for Research on the Wisconsin Economy
Center for Wisconsin State, Local, and Tribal Governance

**The La Follette School has cooperative relationships
with these units at the University of Wisconsin-Madison:**

Center for World Affairs and the Global Economy
Department of Urban and Regional Planning
Nelson Institute for Environmental Studies
Institute for Research on Poverty
International Institute
Law School
Wisconsin Center for Education Research

The *La Follette Policy Report* is a semiannual publication of the Robert M. La Follette School of Public Affairs. The school is a nonpartisan teaching and research department of the College of Letters and Science at the University of Wisconsin-Madison. The school takes no stand on policy issues; opinions expressed in these pages reflect the views of individual researchers and authors.



© 2006 Board of Regents of the University of Wisconsin System and the Robert M. La Follette School of Public Affairs

Welfare and Child Support Policy: Who Should Benefit When Child Support is Paid to TANF Families?

Maria Cancian and Daniel R. Meyer

Maria Cancian is a professor with the La Follette School of Public Affairs and the School of Social Work at the University of Wisconsin-Madison, where she is also director of the Institute for Research on Poverty. Daniel R. Meyer is a professor with the School of Social Work at the University of Wisconsin-Madison. They are the principal investigators for the institute's Child Support Demonstration Evaluation, which is ending after almost nine years of examining the effects of Wisconsin's welfare reform on child support. Their collaborating researchers are Judi Bartfeld, Tonya Brito, Emma Caspar, Judith Cassetty, Jane Collins, Steven T. Cook, Thomas Corbett, Robert Haveman, Thomas Kaplan, Sandra Magaña, Katherine Magnuson, David Pate, Arthur Reynolds, Gary Sandefur, Nora Cate Schaeffer, Judith Seltzer, Geoffrey Wallace, Barbara Wolfe, and James Ziliak. This article draws on papers produced under the Child Support Demonstration Evaluation. Links to all papers prepared as part of the evaluation can be found at www.irp.wisc.edu/research/childsup/csde/csdepubs.htm.

The landscape of policies providing assistance to single-parent families dramatically changed with the 1996 Personal Responsibility and Work Opportunity Reconciliation Act. The long-standing Aid to Families with Dependent Children was eliminated and replaced with Temporary Assistance for Needy Families (TANF). Under TANF, states had considerable flexibility and responsibility for designing new programs. One aspect of the new flexibility was that states were given freedom to change the interface between the child support and welfare systems. States could determine how to distribute child support collected on behalf of children whose mothers were TANF participants.

Most states decided to retain all the child support paid and to use these funds to offset a portion of the costs of administering a child support program and/or the costs of welfare payments. Of all the states, only Wisconsin decided to allow all child support that non-custodial parents paid to pass through to the family and then to disregard those payments in calculating welfare benefits. This policy was put in place in 1997 as part of the original TANF program in Wisconsin, Wisconsin Works.

From its inception, the Wisconsin policy was the subject of a full evaluation. The Child Support Demonstration Evaluation, conducted at the Institute for Research on Poverty at the University of Wisconsin-Madison, included several primary components: a statewide random-assignment experimental evaluation; quantitative nonexperimental evaluations using Wisconsin and national data on child support policies; analyses of policy implementation and monitoring; and qualitative explorations of family dynamics and responses to the new state policies. This article primarily reviews results of the experimental evaluation.

The Policy Context

Prior to 1996, under the Aid to Families with Dependent Children (AFDC) rules, most states provided, or passed through, to families the first \$50 of child support collected each month and disregarded this amount in calculating AFDC benefits. Any child support above \$50 was, however, retained and divided between the state and federal governments. The new program, TANF, allowed each state to set its own policies regarding the state's share of child support payments, although the federal government continued to retain its share. Under TANF rules, most states chose to retain the entire amount of child support collected, passing none of the money through to the custodial parent and children. Two states did otherwise. Minnesota elected a full pass-through, to give the entire amount of support paid to TANF families, but then to adjust the TANF check dollar for dollar (a zero disregard). This approach provides full information to TANF families about how much the noncustodial parent pays in child support, but leaves families with the same income package they would have whether child support was paid or not. Wisconsin devised a different policy, combining a full pass-through with a full disregard. In Wisconsin, each family received the entire child support payment, which the state then ignored in the calculation of TANF benefits. This enables families to see how much is being paid and gives them more income when the noncustodial parent pays support.

Potential Effects of a Full Pass-Through and Disregard

Whether a state disregards child support when calculating benefits essentially determines whether child support will benefit families or taxpayers. Thus, the

clearest effects of a full disregard (compared to a partial disregard or zero disregard) should be that custodial families receive more child support and the government (taxpayers) retains less and thus has higher costs. Because the effects of the policy change flow mechanically, they do not require individuals to change behavior. Thus the primary domains of interest are the effects of the policy on child support received and on governmental costs.

Yet the policy could have other effects, two of which we highlight: whether fathers pay support and whether paternity is established for nonmarital children. Previous child support research has used a model in which the amount of child support fathers pay is related to factors that include noncustodial parents' ability to pay support, their willingness to pay support, and the policy environment. This model predicts that noncustodial fathers who face a policy in which their payments would not fully benefit their children would be less willing (and therefore less likely) to pay support. Ethnographic research suggests that fathers of children receiving welfare realize that any child support paid through the formal system does not fully benefit their children, and therefore they are reluctant to cooperate.

Disregard policy can also affect the extent to which custodial families cooperate with the child support system, which could affect paternity establishment. Research has indicated that in the prior partial-disregard system, some parents strategically collaborated. That is, in exchange for the custodial parent not providing the state information on the noncustodial parent, the noncustodial parent agreed to pay child support informally, which allowed the custodial parent to keep all child support paid. Thus, one effect of a full disregard may be that custodial mothers cooperate more with the child support enforcement system, which could lead to increases in paternity establishment and/or increases in formal child support orders and payments.

In addition to these primary effects, the change in policy could affect mothers' and fathers' employment, earnings, and participation in

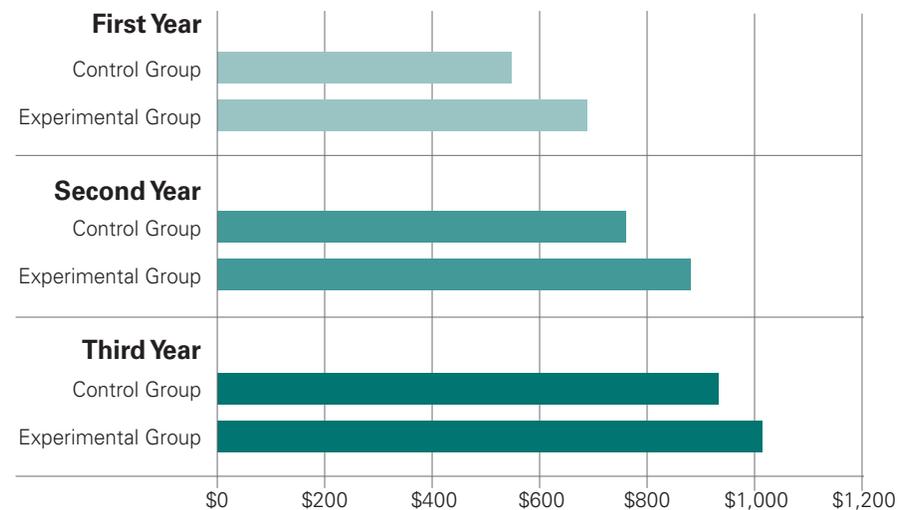
benefit programs; parents' interactions with each other; the noncustodial father's relationship with the children; and the well-being of children.

Child support disregard policy also could affect government costs (the costs to taxpayers). Obviously, giving up child support retentions is a cost. But fully disregarding support could cause custodial families to change other behaviors, which could affect the costs of TANF and other governmental programs. In addition, a simpler system in which all payments go directly to the custodial parent, rather than being divided at the end of each month, could lead to lower administrative costs in the child support system. The ultimate fiscal implications of a policy to disregard all child support will depend on the extent to which increases in cooperation with the child support enforcement system and changes in employment and program participation compensate for the loss in revenue previously collected from child support payments to families receiving public support.

The evaluation was designed to pay particular attention to the way impacts on child support paid and received, paternity establishment, government costs, and other secondary effects changed over time. We anticipated that some effects would grow because new policies often take time to be understood and to have effects. On the other hand, the difference in disregards only affects families when

Figure 1: Average Amount of Child Support Received

The amount of child support received by mothers in the experimental group was greater each year.



The average dollars noncustodial fathers paid each year.

Note: The difference between the experimental group and the control group is statistically significant at the 0.01 level in each year.

they receive Wisconsin Works benefits; once families leave the program, they receive the full amount of support whether they are in the experimental or control group. This could lead to declines in effects over time.

Child Support Demonstration Evaluation

The Child Support Demonstration Evaluation (CSDE) of the Wisconsin policy began shortly after Wisconsin's TANF program, Wisconsin Works (W-2), went into operation. The evaluation's central component is a random-assignment evaluation; most W-2 families received a full pass-through and full disregard of monthly child support, but some child support was withheld from a randomly selected control group. (These families received up to \$50 per month, or 41 percent of what was paid, whichever was greater.) This approach allowed evaluators to attribute any observed differences in outcomes between the two groups to the difference in the treatment of child support. The experimental evaluation and related research drew on a large, longitudinal database incorporating administrative data from several sources, and three waves of data from a longitudinal survey, the Survey of Wisconsin Works Families. The CSDE was completed in several phases that largely corresponded to experiment and policy changes.

Phase 1: The Primary Experimental Results

In the first years of the evaluation, the Wisconsin policy had a variety of positive effects. Most importantly, it increased child support collections and the rate at which paternity was established for children born to unmarried parents, generally without significantly increasing government costs.

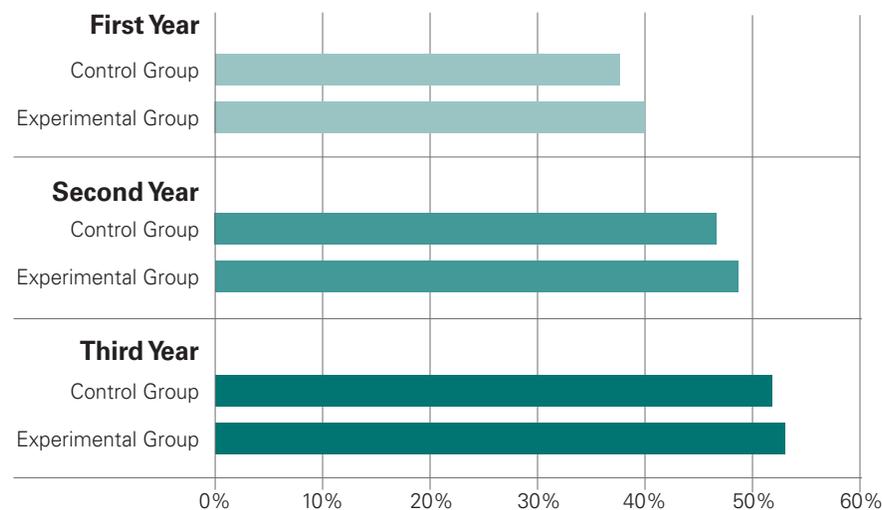
Figure 1 examines the amount of child support received by the nearly 16,000 custodial mothers who entered W-2 in its first nine months. For both the experimental and control groups, the amount of support increased as time passed. The experimental group received more child support in each year, with a difference of \$138 in the first year. The impact was larger among some subgroups, including those that had a child support order at entry (a difference of \$190).

This impact largely reflects the mechanical effects of the policy change, in which experimental-group mothers receive the full amount paid and control-group mothers do not receive the full amount if they receive W-2 benefits. We also find small, but statistically significant differences in the percentage of fathers who pay support, and thus the percentage of mothers who receive support. As shown in Figure 2, the proportion of mothers receiving support increases over time. Those in the full disregard group are more likely to receive support in each year, but by the third year the difference is no longer statistically significant at conventional levels. Among those mothers who had not received AFDC in the prior two years, the impact was particularly large, with 35 percent of experimental-group mothers receiving support in the first year, compared to 30 percent of those in the control group.

We also find significantly higher rates of paternity establishment in the first year. Among the nearly 15,000 non-marital children who did not have legal fathers when they entered W-2, 17.2 percent had paternity established by the end of the first year, compared to 15.7 percent of those in the control group, as shown in Figure 3. The difference disappeared by the end of the second year. Similar to the other outcomes, the difference is particularly large for those who had not

Figure 2: Percentage of Mothers Receiving Child Support

The percentage of mothers receiving child support was greater in the experimental group for all three years.



The percentage of mothers receiving child support.

Note: The difference between the experimental group and the control group is statistically significant at the 0.05 level in the first year and at the 0.10 level in the second year.

recently received AFDC. In summary, a full disregard is associated with faster paternity establishment, but not with a long-term increase.

While there are differences in some components of government costs, we find no difference in overall government costs between the experimental and control groups. Although more child support is passed through to those in the experimental group, not all of this is at the expense of government, since some consists of additional support that would not have been paid in the absence of the full disregard. More important, the reform generated cost savings in other areas, especially in W-2 cash payments. Finally, our estimates of cost differences do not consider the administrative savings that may have resulted from a simpler child support system.

Phase 2: Longer-Term Outcomes for Families

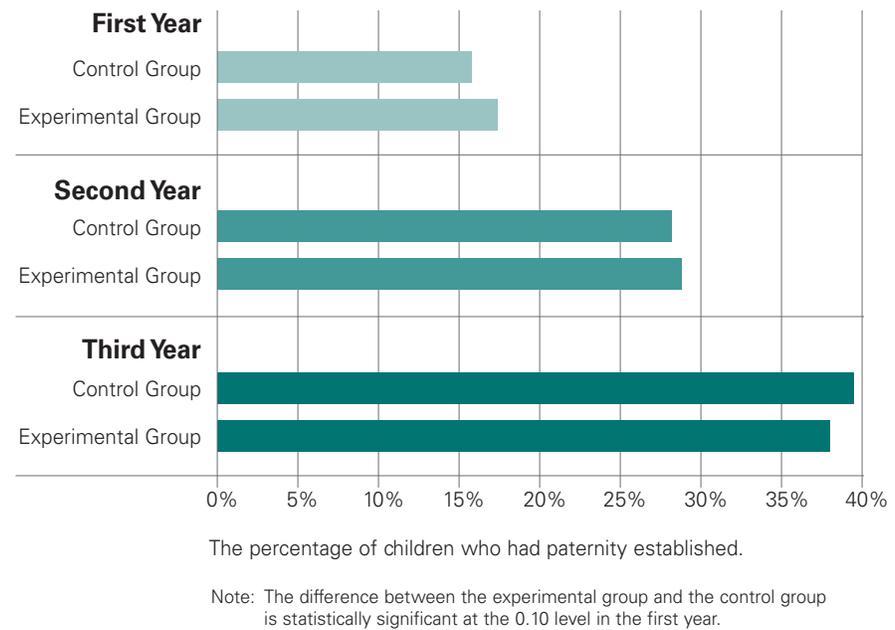
In a 2003 report, we present longer-term outcomes for two groups: the first cohort of families (those in the Phase 1 report), and a second cohort of randomly assigned families that entered W-2 six months or more after the first group. The report covered three to four years of experience with the new policy. Evaluation continued to demonstrate that the pass-through and disregard policy increased child support receipts. Because the later cohort entered after some improvements in W-2 and child support pass-through policy implementation, we had hypothesized that there would be a stronger experimental effect within this group. For the most part, we did not see that effect.

The End of Random Assignment

Beginning July 1999, all new families entering W-2 received the full pass-through and disregard. Among families already receiving W-2, those cases previously assigned to the control group continued to receive a partial disregard until July 2002. At that point all cases, new and old, began to receive the full pass-through and disregard. These changes defined two additional cohorts of families receiving the full disregard—those entering during a time of transition,

Figure 3: Percentage of Children for Whom Paternity is Established

The percentage of children with paternity newly established was greater in the experimental group in the first two years.



when some older cases still received a partial disregard, and those entering during a time when all cases received the full disregard.

Comparisons of Outcomes in the Original and Extended Evaluations

A final annual report, completed as part of the extended evaluation, describes findings from six years of observations for the original two randomly assigned cohorts of families, and two to five years of observations for the two later-entering cohorts that received the full disregard. The report corroborates results from the earlier years of the experiment that showed increased paternity establishment and child support payments. It also shows that most of the experimental effects did not persist, because families moved quickly off W-2 cash assistance and thus were only briefly exposed to differences in the experimental treatments.

The finding of limited longer-range effects does not necessarily mean that the full pass-through and disregard policy has not been beneficial. The narrow differences separating the partial disregard as defined in the evaluation and the full disregard received by most families mean that all W-2 cases in Wisconsin have benefited from a fairly generous disregard policy, compared to the no-disregard policy the majority of states use.

Nonexperimental Evaluations: Contexts for the Experimental Evaluation

Experimental and nonexperimental evaluations have different strengths. A key limitation of experimental designs are that they provide information only on the comparison of the policy regimes actually tested; they tell us little about the effects of other potential policies. Currently, most states retain all child support, and in those that disregard some amount, the typical amount is \$50 per month. Thus the Wisconsin experimental design results, which rely on a comparison of a full pass-through/disregard with a pass-through/disregard of the greater of \$50 or 41 percent, cannot provide direct information on the effect of a full pass-through/disregard compared to the state retaining all child support, or compared to a straight \$50-per-month pass-through/disregard.

In related research we compare outcomes as disregard policies varied across states and over time. In theory, this approach should enable us to estimate the effects of the full range of disregard policies observed. The key limitation to this analysis is the possibility that, even after we control for other relevant factors, the groups facing different regimes differ in ways other than the disregard policy. With these trade-offs in mind, we summarize the results of a nonexperimental analysis.

Disregard policies have changed over time—from typical regimes of no disregard, to a \$50 disregard, to a state option of up to \$50. Policies have also varied from state to state—some states had fill-the-gap policies, some obtained waivers to change policies before TANF came into existence—and TANF brought considerable policy variation across states. We examined disregard policy for the AFDC/TANF caseload for each state and the District of Columbia for 1985-1998. We considered the proportion of cases with paternity established, the proportion of cases with collections, and the amount collected.

We found that a larger disregard is associated with a statistically significant increase in the rate of paternity establishment and a small statistically significant increase in the proportion of cases with collections. The size of the disregard does not have a statistically significant relationship with the average collection per case among cases with collections. Taken as a whole, the results support the conclusion that increasing the disregard will not only increase the receipt of child support (a mechanical effect) but will also increase the probability of payment of child support (a behavioral effect). The confirmation of this key result from the CSDE experiment in the non-

experimental study relying on national data is encouraging. Neither the CSDE nor the nonexperimental study shows a strong relationship between the disregard and the amounts of child support paid. Since a higher disregard is associated with a greater proportion of noncustodial parents paying support, it may be that lower-income noncustodial parents are overrepresented among those entering the system. Finally, the CSDE experiment suggests that paternity establishment proceeds more quickly for children eligible for a full disregard, although the difference in paternity rates for the sample as a whole disappears after the first year.

Broadening the Scope of CSDE Research

In addition to the experimental and nonexperimental evaluations, the CSDE included a range of related research drawing from the core administrative and survey data, from ethnographic and field research, and from specialized data collection. For example, an examination of participants' knowledge of child support policy rules shows that many parents do not fully understand the policy. This lack of understanding may have limited the policy's effectiveness.

Our analysis of the effects of the full pass-through and disregard on marriage, cohabitation, and living arrangements finds evidence that mothers receiving the full pass-through and disregard were less likely to be cohabiting with men who were not the fathers of their children and were more likely to be single. Reports from this research are available at www.irp.wisc.edu/research/child-sup/csde/csdepubs.htm.

What the Future Holds

Wisconsin's child support disregard policy changed again in January 2006; the federal waiver that permitted the CSDE has expired; and the full-pass through is being phased out. This most recent change falls outside the scope of the CSDE. At the same time, federal policy is changing in ways to encourage other states to follow Wisconsin's innovative original approach. The 2006 TANF reauthorization allows states to pass through and disregard the first \$100 per month of child support for one-child families and \$200 per month for larger families without reimbursing the federal government for its share of the support. The results of the CSDE demonstrate the potential advantages of this policy. ♦

program but who are simply put on a waiting list for a year. In this case, both the treatment and control groups experience the same anticipation of participation, so the anticipation effect is effectively controlled. This allows researchers to be more confident that their analysis is restricted to the intended program features.

With regard to program design, one potential concern is that since many of the implemented and evaluated childhood interventions are targeted toward disadvantaged populations, their overrepresentation may bias the results. This may be the case when viewing the results of an evaluation in the context of a group whose composition differs from those in the evaluation or from the population as a whole. However, the results are valid for the population that the sample represents. So, as long as policy decisions are made with this in mind, there should be no difficulties in extending a childhood intervention to similar populations. With the variety of reasons to focus programs on disadvantaged populations — including equity and the possibility of very high rates of return — we do not view the abundance of programs with a limited scope as a problem.

As for sample collection, size is probably the most important consideration. In general, the larger the sample, the better, since a larger sample yields greater statistical power and allows researchers to estimate the level of program effects with higher precision. With small samples, it is often difficult to reject the hypothesis that a childhood intervention had no effect on the population, even if there were effects.

Related to sample size is sample attrition. In evaluations, researchers saw a range of success in eliciting responses in follow-up interviews. For example, they may have had trouble contacting people who moved or otherwise became unreachable. Other reasons that could be more damaging to the evaluations may include participants who intentionally avoided the interviews or who sought to be interviewed. In general, we viewed smaller attrition rates as unambiguously better, and so ranked those studies with lower attrition rates as more favorable. Additionally, we accounted for situations in which researchers attempted to show that attrition did not affect the sample composition for important characteristics, or in which researchers suspected that attrition affected

the sample in specifically predictable ways.

Attrition is only one of the ways in which sample selection may be introduced into an evaluation, so we are also cautious of other sample selection possibilities when considering childhood interventions. For example, there may be sample selection, on the part of the program designer or the participants, in either the initial recruiting process or the subsequent baseline evaluation, or both.

Next, we take into account special considerations that are important when thinking in the context of economic growth. Extended follow-up periods are especially important. Many of the economic returns to childhood interventions and, especially,

early childhood interventions occur relatively far into the future. For example, major determinants of economic growth include the labor supply decisions of workers as well as the rate of technological innovation, both of which most often occur in adulthood. In the case of some early childhood interventions, these returns may only be directly measurable up to 20 years after the programs are implemented. So, economists rely on techniques to forecast specific program effects to then estimate the benefits or effects on economic growth.

Another consideration that is extremely important is the idea of outcome relevance. Some outcomes are much more feasibly descriptive of human or health capital than others. For example, cognitive test scores are often viewed as at least a somewhat reasonable proxy for human capital since they capture both heritable features and investment in a child. The relative values of social or psychological outcomes are more difficult to evaluate in the current context because they do not as readily fit into existing economic frameworks. At the least, we view outcomes that are more directly linked to human or health capital development as more conducive to economic growth research than more indirect outcomes, since the latter may introduce error.

Finally, we consider any general features of the childhood intervention that may influence whether it can be replicated. For example, results from a study that is restricted to one location may not be as generalizable as studies performed in multiple locations. More specifically, the location may have idiosyncratic characteristics that can introduce special effects into

With small samples, it is often difficult to reject the hypothesis that a childhood intervention had no effect on the population, even if there were effects.

the evaluation. Another example of this is person-specific effects. A childhood intervention evaluation that consists of a reasonable sample size but nonetheless depends on the specific characteristics of a few individuals (such as home visiting nurses) may warrant caution.

Most Promising Childhood Interventions

We begin this section with a discussion of a selection of programs that we believe to be well suited to further research in the context of economic growth, including further evaluation and forecasting via economic models. Using the criteria discussed in the last section, we studied about 200 childhood intervention programs implemented through various means and then evaluated in experimental or quasi-experimental settings.

Table 1 summarizes some of the most promising programs, including two we do not discuss in detail: Career Academics and the Job Corps. Both appear very promising as interventions to encourage work and increase the productivity of young adults.

Table 2, available online at www.lafollette.wisc.edu/publications/policyreports/wolfetable2.pdf, lists other promising programs.

Nurse Family Partnership

To improve prenatal care and the health and caretaking of infants, the Nurse Family Partnership assigns nurses to visit the homes of disadvantaged women who are new mothers or pregnant. The program has been implemented for evaluation at multiple sites, including Memphis, Tennessee; Elmira, New York; and Denver, Colorado. Because a long follow-up duration is important in the context of economic growth, we focus on the Elmira site because it offers the opportunity to view outcomes after a 15-year follow-up. Here, we review the results from the Elmira site reported in 1998 by D. Olds, P. Hill, and E. Rumsey.

In collecting sample data, the researchers recruited women who were less than 25 weeks pregnant and who had no previous live births. Most of the women were young, married, or of low socioeconomic status. Between April 1978 and September 1980, 400 women were enrolled in the program out of the 500 invited to participate.

After a baseline interview, the women were randomly assigned to one of four treatment groups

arrayed by intensity of intervention. The first group of children was screened at 1 and 2 years and referred for care only if necessary. The second group received free transportation for medical care until the child was 2 years old. The third group received nurse home visits during pregnancy in addition to transportation. The fourth group received visits from nurses until the child was 2 years old plus transportation.

Children were assessed at different stages of their development, including a 15-year follow-up. This is a substantial follow-up duration with results that are likely to last. The relatively low attrition rate of 22 percent is also a strength. Children whose families received visiting nurses reported fewer arrests, convictions, and violations of probation, but they reported police stopping them more frequently than did those children

whose families did not receive visiting nurses. Minor crimes accounted for most of the program's effect. Children whose families received visiting nurses reported fewer sexual partners, lower rates of cigarette smoking, and fewer days of alcohol consumption. Results for illegal drug use and teacher reports of problem behavior did not follow the expected pattern. Families with low socioeconomic status showed the greatest effects from the treatment across most variables.

In drawing conclusions from the Elmira study some caution is appropriate. Although effects were strong even after 15 years, the results were driven mostly by the low socioeconomic status group. Results are from a rural area (in part, addressing this is the motivation for studying the program's effects at other sites.) Idiosyncratic characteristics of the nurses may contribute somewhat to the effects.

Chicago Child-Parent Center Program

Aside from Head Start, the Chicago Child-Parent Center program (CPC) is the oldest federally funded preschool program in the country. The program was started in 1965 with the goals of improving the academic achievement of disadvantaged children and involving parents more in their children's education. CPC was a center-based intervention that included education, family, and health services to all participants. Children entered CPC between ages 3 and 4, and services were usually offered until age 6, but they

The relative values of social or psychological outcomes are more difficult to evaluate in the current context because they do not as readily fit into existing economic frameworks.

Table 1: Most Promising Childhood Intervention Programs

Program	Category	Description	Sample
Nurse Family Partnership	Home Visiting	Program provides nurse home visits and referral services for mothers and children until child is 2 years old. Study with longest measurement duration was administered in Elmira, New York, starting in 1978. Women were recruited if they had no previous live births; most were young, unmarried, or of low socioeconomic status.	500 women were recruited, and 400 enrolled. Participants were randomized into four treatment groups. Attrition rate was 22 percent.
Chicago Child-Parent Center program	Preschool	Nonrandomized group of children who lived in low-income areas were invited to participate in a program that included education, family, and health services. Program began in 1965 with children at ages 3 to 4; included services to age 6 with some continuing to age 9. The matched group lived in other low-income areas of Chicago where CPC was not offered and primarily consisted of full-day kindergarten.	989 participants. 550 participated in other related but less ambitious programs.
Class Wide Peer Tutoring	School	Participants are recruited within classrooms for a peer tutoring process with rewards for progress. Those included in the evaluations were from grades 1-6 in inner city low-income neighborhoods in Kansas City, Kansas. The program has also been tested with children with learning disabilities and children in higher income areas.	Sample size groups for the evaluations ranged from four to more than 400.
Healthy Kids	Health	Healthy Kids provides health insurance coverage to children in Santa Clara County, California, with household income below 300 percent of the federal poverty line and who are ineligible for the major state health insurance programs. More than 80 percent of the population served is Latino. The data for the analysis is drawn from a survey of program participants over a single year. Recent enrollees are used as a comparison group to measure the program's effects.	Total sample size was 1,235, and 89 percent of enrollees contacted for the survey responded.
Big Brothers Big Sisters of America	Mentoring	This is a nationwide mentor program that matches adult volunteers with children who tend to be from single-parent families. The study population included 959 children age 10 to 16 who applied to the program in 1992 and 1993; half were assigned to the treatment group while half were assigned to a waiting list.	The study chose 1,138 youth to be in the study; 84.3 percent completed both surveys. 378 of the 487 in the treatment group were matched. More than half were minorities; almost all lived with one parent.
Career Academies	Job Training	This program seeks to keep students in high school and prepare them for the transition to further education and employment. The study was conducted at nine high schools in urban, low-income areas nationwide. Participants were selected for the study by lottery starting in 1993.	1,764 students completed the program and 1,458 (83 percent) completed the follow-up survey. 55 percent finished the program, with varied dropout rates.
Job Corps	Job Training	For 13 months in 1994-95, applicants nationwide were randomly assigned to two groups, one that could enroll in Job Corps and another that could not, though they could receive other services. Most applicants were high school dropouts, most were minorities, and about a quarter had arrest records. The program offered residential and nonresidential settings. It provided education and vocational services, counseling, and medical care.	9,400 were in the treatment group and 6,000 in the control group. Follow-up interviews had an 80 percent response rate.

Table 1: Most Promising Childhood Intervention Programs (continued)

Outcomes	Follow-Up	Replicability	Relevance	Program
Participants reported fewer arrests, convictions, violations of probation, lifetime sex partners, cigarettes smoked per day, and days of alcohol consumption in the previous six months. Effects greatest among mothers with low socioeconomic status.	15 years.	There may be issues with respect to the specific implementation processes by nurses.	Relevant. There is a large number of significant outcomes, and the follow-up duration extends well into high school.	Nurse Family Partnership
Outcomes included grades and other achievement measures, parental involvement, arrests, grade retention, special education, and completion of high school and post-secondary education. In nearly every case, participants performed significantly better than those in the comparison group.	Ongoing, and this study evaluates through age 23.	Likely to be replicable, but a random design would have increased confidence.	Very relevant. Measurement duration extends into the adult wage earning period, and there are many measures on which to base further analysis.	Chicago Child-Parent Center program
Students who started the program in first grade performed significantly better on national standardized tests for math and reading. Other outcome measures included time spent on task, grades received, and other standardized test outcomes.	Evaluators claim that results persist for at least three years.	Appears to be easily replicable and is of very low cost. The main issue of replicability is in establishing the pairing mechanism.	Relevant. Measurement outcomes are evaluated only a few years past elementary school, but the measured variables are important.	Class Wide Peer Tutoring
Results show that program reduced unmet need, improved access to health care, and boosts health-care usage.	Effectively one year.	Difficulties in replicability may arise from the fact that the program was administered in a single county. Results may only generalize to similar populations.	Somewhat relevant. The outcomes measured only relate to economic growth through improved health. Measurement duration was short.	Healthy Kids
Findings based on self reports. Those paired were 46 percent less likely to initiate drug use; 27 percent less likely to initiate alcohol use; one-third less likely to hit someone, skipped half as many days of school, and showed modest gains in school achievement.	18 months.	Replicable. The program is in many cities and has a long history.	Somewhat relevant. The outcome measures themselves are tertiary to economic growth, while the measurement duration may in some cases extend close to working age.	Big Brothers, Big Sisters
Men in treatment group earned significantly higher wages. The effect was concentrated among members who were at high or medium risk of dropping out of high school upon enrollment.	Four years after the participant's scheduled graduation from high school.	The study was implemented at sites that had operated Career Academies for two years. This may have introduced some selection bias.	Very relevant. Wage data are clearly important for economic growth analysis.	Career Academies
Outcome measures included hours of additional education, GED, literacy tests scores, earnings, cash welfare assistance, and arrests. Significant differences in crime accounted for substantial share of program benefits.	Outcomes measured at 12, 30, and 48 months.	Authors say this nationwide program had similar impact across sites.	Very relevant. Measurement duration extends into adult wage earning period, and there are many measures on which to base further studies.	Job Corps

sometimes continued until age 9. For the preschool portion of the program, children attended for half-day sessions during the nine-month school year. Beginning with kindergarten, the program consisted of full-day participation during the school year. Key features included smaller-than-normal class sizes, with one to eight children in preschool classrooms and one to 12 children in kindergarten classrooms. Also, parents were encouraged to directly participate in the education process. Finally, children who participated in CPC were drawn from low-income, largely African American areas of Chicago.

In the 1999 Chicago Longitudinal Study, Arthur J. Reynolds describes an evaluation of the effects of CPC on a number of outcomes. The study includes all 1,150 children who participated in the kindergarten segment of the 20 CPC in operation between 1985 and 1986. Since most of the participants were from low-income areas, a matched comparison group consisted of 550 students from five randomly chosen Chicago area public schools that served similarly disadvantaged children. This nonrandom assignment process may have introduced selection bias into the subsequent statistical analysis. For example, among low-income areas there may have been competition to live in districts that CPC served if people believed that CPC was more successful at improving student achievement. Insofar as any important differences between the CPC group and the comparison group are controlled for by observable characteristics, however, there will be no selection bias.

Researchers evaluated 772 CPC group participants and 392 comparison group participants at age 14 for a range of outcomes. This represents an overall attrition rate of less than 25 percent, a relatively low rate considering the follow-up duration. Indeed, the favorable combination of attrition rate and follow-up duration adds confidence that program effects will persist. In addition, evaluations of this sample are ongoing, so there is promise that outcome measurements will extend well into participants' wage earning years. Evaluations to date include higher school achievement, lower dropout rates, fewer events of juvenile delinquency, and higher probability of attending a college or university. Many of these are directly relevant to economic growth. With appropriate caution due to the nonrandom design of the study, we believe that the program results warrant further study.

Class Wide Peer Tutoring

The Class Wide Peer Tutoring program (CWPT) was created to engage elementary-school children in classroom activities and to foster greater student achievement. Program evaluators C.R. Greenwood, J.C. Delquadri, and R.V. Hall reported in 1989 that development of the program was motivated by the fact that children with more of an "opportunity to respond" in class tend to be more successful in school. As a result, the program seeks to maximize engagement by pairing

students in a tutor-tutee relationship where students administer questions to each other. By identifying and rewarding daily and weekly winners in a competition among groups, the children are motivated to perform well.

Researchers conducted a longitudinal evaluation of students between the first and fourth grades with a random experimental design in which a treatment group was drawn from children in families with low socioeconomic status. Those not selected served as the control group. In addition, the researchers identified a group of students in families with high socioeconomic status to serve as a comparison group. Children attending nine schools in Kansas City, Kansas, form the sample.

Baseline evaluations were given at the outset of the program (at the beginning of first grade) and follow-up evaluations were given after students completed fourth grade. Attrition is an issue, given a 56 percent attrition rate, including a 68 percent rate for the experimental group. Most of this is due to reassigning children to schools not participating after the district closed one school. Outcome measures included standardized achievement tests relative to national averages. The researchers found significant differences between the treatment group and the control group, and between the comparison group and the control group. On average, the treatment group scored more than 10 percentage points higher than the control group on the achievement test.

The relationship between the program and economic growth models is somewhat promising. Achievement test scores provide a reasonable link to the human capital framework, and the measurement duration of four years is a significant period of time. However, the children were only last evaluated in the fourth grade, well before impacts on economic growth might be realized. The location in one city creates some question of whether results could be replicated elsewhere.

The Healthy Kids program reduced the proportion of children with unmet need in the past six months from 22 percent to 10 percent across all categories.

Healthy Kids

To address gaps in health insurance coverage among low-income families, the Healthy Kids program was started in Santa Clara County, California, in 2001. The Healthy Kids program provides health, dental, and vision insurance coverage for children in families that earn below 300 percent of the federal poverty line and are not eligible for the two state-funded programs. Since citizens below 250 percent of the federal poverty line are eligible for at least one of the state funded programs, the majority of the participants in Healthy Kids are not otherwise covered because of their immigration classification. The program has covered more than 30,000 families since its inception.

More than 80 percent of the participants are Latino and live in non-English speaking households, most of the children live in two parent households, and the vast majority live in households with at least one working parent, suggesting that most of the participants' parents work in jobs that do not offer health benefits. Prior to the program, 63 percent of the children had no health insurance within the six months before enrollment, and 45 percent were never covered.

A survey-based evaluation was not a random assignment study, but instead compared two groups of participants separated by the timing of coverage, C.A. Trenholm, E. Howell, D. Hughes, and S. Orzol report in a 2005 study. "Established enrollee" children were involved in the program for approximately one year and had completed renewal of coverage (treatment group), while "recent enrollee" children were those interviewed when they became eligible for Healthy Kids, but who had not yet participated (control group). Insofar as the sample was a random selection of participants from each of these groups and as long as there were no significant idiosyncratic year shocks, the design will yield reasonable estimates of the program effects. The total number of respondents was 1,235, the survey response rate was high (89 percent).

In terms of outcomes, the program reduced the proportion of children with unmet need in the past six months from 22 percent to 10 percent across all categories, including well-child visits, sick-child visits, specialty care, prescription medications, and

dental care. There were also substantial increases in the proportion of children who had a usual source of primary care (50 percent to 89 percent) and those with a usual source of dental care (29 percent to 81 percent.) There are some caveats in considering the program in the context of economic growth: the population is a specific, though large and growing, group whose results may not generalize to other areas or groups; relating some outcomes to economic growth may be difficult, as they are only measured over one

year; and finally, the outcome measures themselves are only secondary to variables directly relevant to economic growth, so the introduction of error in tying them into economic growth models is a concern.

Big Brothers Big Sisters of America

Big Brothers Big Sisters (BBBS) is the oldest and one of the most well-

known mentoring programs in operation. At any given time, there are approximately 75,000 mentor-child matches nationwide. Local programs function as affiliates of the national program, which plays the important role of providing criteria and standards for the appropriate functioning of the local programs and development of matches. This consistency across programs is important if any experimental findings are to be generalized.

For the most part, the programs recruit and screen youth and mentors who have an interest in forming a match. Then, based on background characteristics and stated preferences, the programs match youth and mentors. Mostly, the children are from disadvantaged families, often with a single parent. The matched youth and mentor will then meet approximately three or four times each month for at least a year.

A study selected eight BBBS programs nationwide for evaluations based on program size and geographic diversity, J.B. Grossman and J.P. Tierne reported in 1995. The sample consisted of 1,138 youth between the ages of 10 and 16 (most between 10 and 14) most of whom were from disadvantaged families. After baseline interviews were completed, the youth were randomly assigned to the treatment or control group where members of the treatment group were immediately matched to mentors while

This review demonstrates the importance of including a careful plan to evaluate childhood intervention programs. Policymakers and program developers with creative ideas for childhood interventions may overlook evaluation or fail to allocate the resources necessary for establishing their success.

those in the control group were placed on an 18-month waiting list for the program. Placement on a waiting list is a useful control for any effects that might be associated with an individual's anticipation of participating in a program. After 18 months, each participant was administered a follow-up interview with a low 15.7 percent attrition rate. In terms of outcomes, participants were 46 percent less likely to initiate alcohol use, were one-third less likely to hit someone, skipped half as many days of school, and showed modest gains in school achievement.

Summary/Conclusions

The five programs discussed here, plus the two additional programs outlined in Table 1, are interventions that appear likely to increase economic growth. This review also demonstrates the importance of including a careful plan to evaluate these childhood intervention programs. Policymakers and program developers with creative ideas for childhood interventions may overlook evaluation or fail to allocate the resources necessary for establishing their success. When evaluation is included in program design, it often is short-term in nature, which impedes linkage of the intervention to economic growth.

Evaluations ideally should include random samples of sufficient size, with treatment and control groups defined so that differences can be fairly attributed to the effects of the childhood intervention program. Also, random samples can be crafted from neighborhoods or waiting lists rather than individuals if that sort of design is easier for evaluators to establish. Whether program results can be replicated with other populations also is important. Follow-up with program participants, with as little attrition as possible, must take place over a long time period to account for contributions to economic growth. And, the relevance of the study's outcome to economic growth must be considered.

Our research highlights childhood interventions that are likely to lead to increases in economic growth and outlines criteria program developers and researchers can use to evaluate these programs and to illustrate their efficacy in improving human and health capital for our entire society. ♦

Director's Perspective *continued from page 1*

Journal and The Economist. A number of La Follette faculty and students are studying this issue, especially as it relates to the University of Wisconsin-Madison.

In addition, La Follette is proud to be host of the 2006 meeting of the Association of Public Policy and Management (APPAM) in Madison in November at the Monona Terrace, the convention center on Lake Monona based on a design by Frank Lloyd Wright.

La Follette faculty play a prominent role in APPAM. David Weimer is APPAM's president; Maria Cancian serves as vice president; and Carolyn Heinrich and Geoffrey Wallace sit on the Policy Council. At the Madison conference, 11 La Follette faculty members are scheduled participants, the same as the number who presented at the 2005 fall conference in Washington, D.C. We all are pleased with the opportunity to share our city and our campus with APPAM conference attendees.

The APPAM meetings also celebrate the 40th anniversary of our sister on-campus institution, the Institute for Research on Poverty, with two conference panels and a reception with all current and former IRP directors (including three faculty now at La Follette) in attendance.

The reciprocal relationship that IRP and the La Follette School share shows in this issue of the Policy Report.

Maria Cancian and co-author Daniel Meyer report on a major project they and many others have worked on for nearly a decade: the State of Wisconsin's Child Support Demonstration. Part of the 1996 welfare reform, this project examined the effects of allowing all of an absentee parent's child support payments to be paid to the resident parent and child, rather than it being retained to offset welfare benefits. In essence, the demonstration permitted the resident family and children, rather than taxpayers, to benefit from additional child support payments.

The paper I co-author with Nathan Tefft emphasizes the importance of comprehensive program evaluation in the context of childhood interventions. After making an argument on the potential importance of childhood interventions, we point out that long-term assessments are essential for us to learn what works (and for whom) in order to maximize the returns on investment in children.

Yin-Wong Cheung, Menzie Chinn, and Eiji Fujii explore several methods of evaluating whether the Chinese currency, the renminbi, is undervalued, and if so, by how much.

As you probably have noted, this issue is the first to appear since I have become La Follette's director. I am pleased to serve in this capacity. I look to all of La Follette's alumni and supporters for help in sustaining the school as one of the nation's premier public policy research and training centers. ♦

The Debate about the Undervaluation of China's Renminbi

Yin-Wong Cheung, Menzie D. Chinn, and Eiji Fujii

Yin-Wong Cheung is on the faculty at the University of California, Santa Cruz and the University of Hong Kong. Menzie D. Chinn is a professor with the La Follette School of Public Affairs at the University of Wisconsin-Madison and a research fellow with the National Bureau of Economic Research. Eiji Fujii is on the faculty of the Graduate School of Systems and Information Engineering at the University of Tsukuba in Japan. This article is based on a paper the authors presented at a conference on "Financial and Currency Integrations" in September 2006 at the University of California at Santa Cruz.

"[T]he Treasury Department is unable to conclude that China's intent has been to manage its exchange rate regime for the purposes of preventing effective balance of payments adjustment or gaining unfair competitive advantage in international trade."

— **Former Treasury Secretary John Snow**

May 10, 2006, press conference



"[The] Chinese currency is grossly undervalued. The steps taken by China last summer and again earlier this week to revalue the yuan, while welcome, can't even be described as baby steps in addressing the problem of China's tight currency controls."

— **Senator Elizabeth Dole (R-N.C.)**

May 18, 2006, press conference.



As these quotes indicate, China—and Chinese economic policy—has loomed large on the global economic stage in recent years. And yet, even as bipartisan bills to impose economic sanctions on China are winding their way through the Congress, tremendous misunderstanding and confusion surrounds what currency undervaluation is and why it might be important.

For economists, a currency is undervalued when the rate at which it exchanges for foreign currency is higher than what economic fundamentals indicate it should. In China's case, the allegation is that it takes more units (yuan) of the Chinese currency (the renminbi) to buy a single dollar than is viewed as appropriate. The quintessential problem is that "undervaluation"—or currency misalignment in general—is in the eye of the beholder. Whether one sees a misalignment depends upon which economic model one has in mind.

In this article, we examine this question, not with a mind to proving that our approach to

determining the fair value of the renminbi is the right one. Rather, we wish to highlight the uncertainty surrounding the issue, both in terms of figuring out which is the right economic model, as well as in implementing the actual mechanics of calculating the degree of undervaluation. Finally, we talk about what it would mean for the renminbi to be revalued.

Why is an Undervalued Renminbi Important?

The most common definition of an appropriately valued exchange rate is one that sets the trade balance at or close to zero. The easiest way to think about this is to consider what the exchange rate does. When the renminbi is strong, say against the U.S. dollar, then, holding everything else constant, only a few renminbi are needed to buy a single U.S. dollar, and more things made in the United States that can be bought with those dollars. People (and firms) responding to this relatively low price will tend to buy lots of American goods and fewer Chinese goods, all else held constant because the U.S. goods are relatively inexpensive. Now reverse this logic: When the renminbi is weak (undervalued) then Chinese goods are relatively cheap and thus in high demand. Using this logic, China's currency is self-evidently undervalued, as most projections are for a trade surplus with the world of \$200 billion in 2006, much of it with the United States. As Figure 1 shows, even as the renminbi has appreciated slightly over the past year, the trade surplus has grown rapidly.

So, potentially, the undervalued renminbi is important in that China runs a trade surplus and the rest of the world a deficit with China. Of course, few in the United States complain when the United States runs a trade surplus (which we last did in the early 1980s), so can this equating of surpluses = good and deficits = bad be correct? Clearly, the answer is no. Some countries should be running surpluses and some should be running deficits,

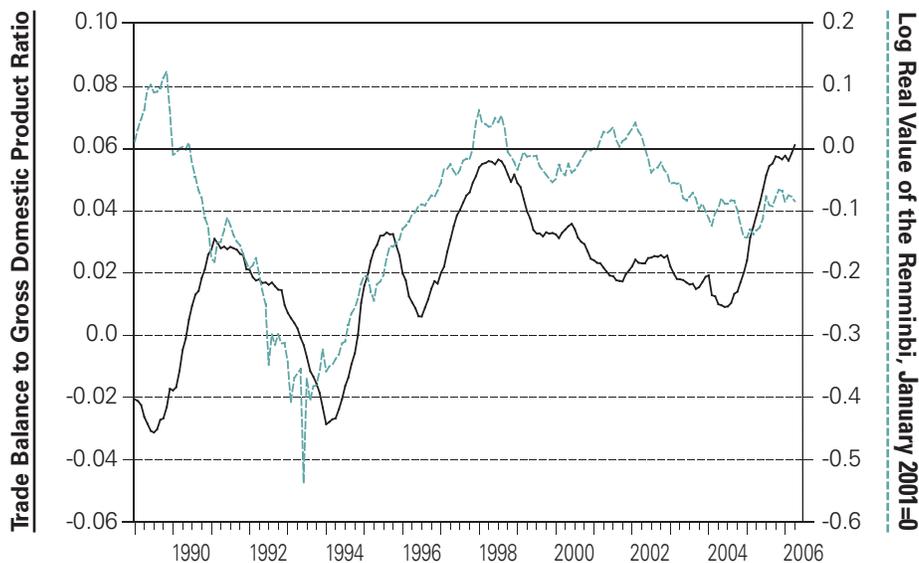
depending upon their situation in the world.

A useful analogy is that of a household. When a household hits a period where income is temporarily low, but prospects for future income are bright, then it might make sense to borrow a bit so as to keep consumption relatively smooth. In doing this, the household is “importing” more than it is “exporting,” running the analog of a trade deficit.

Most observers in policy circles understand this point and make a more nuanced argument—that the trade surplus, net income from abroad and “stable” capital inflows (such as foreign direct investment) should be close to zero. This is sometimes called the “basic balance.” In this view, the appropriate exchange rate is the one that achieves this goal. Yet another approach is to see whether the foreign exchange reserves are rising, i.e., a country is increasing its holdings of other countries’ currencies. The view is that when foreign exchange reserves are rising rapidly to levels that far exceed what would be needed for “a rainy day” (in terms of exports or lending from abroad), then the currency is undervalued. On this count, once again, it looks like the renminbi is undervalued, as Chinese foreign exchange reserves have climbed to new heights, as Figure 2 illustrates.

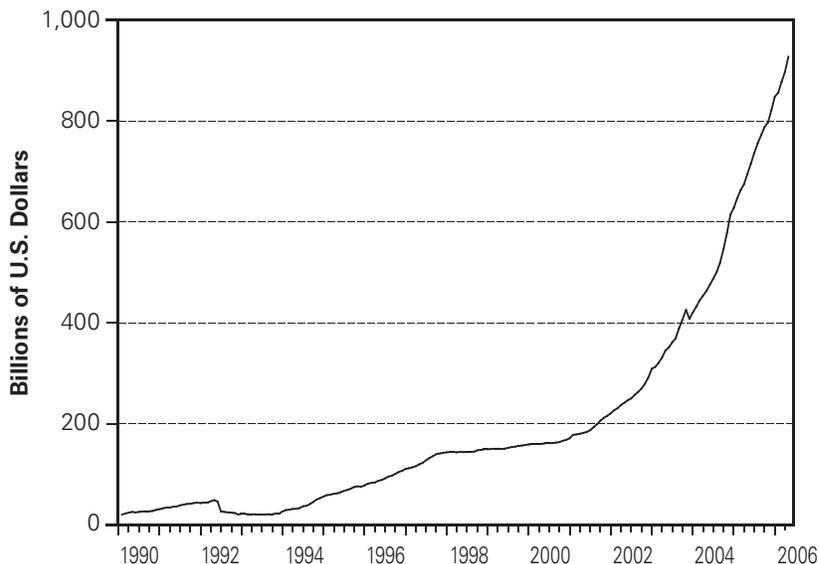
But when is a basic balance “too big”? Or for that matter, when are reserves “too big”? Foreign exchange reserves on the eve of the Mexican peso crisis of 1994 were thought to be so large as to be able to address any possible financial crisis. Before the 1997 East Asian crises, many thought the same applied to Korea and Thailand. So one should be a little circumspect about this way of viewing overvaluation, even if one admits that the approach seems common-sensical.

Figure 1: China’s Trade Balance and Value of Renminbi



Sources: International Monetary Fund’s International Financial Statistics; International Monetary Fund’s World Economic Outlook for April 2006; and authors’ calculations.

Figure 2: China’s Foreign Exchange Reserves



Note: These foreign exchange reserves exclude gold.
Source: International Monetary Fund’s International Financial Statistics

We argue in a forthcoming paper for the use of other approaches to augment our understanding of the renminbi’s equilibrium value. The two main approaches we focus on are the productivity model and a framework that views differential price levels across countries as being functions of per-capita income levels.

Productivity and Relative Income

The role of productivity is central to thinking about the evolution of the Chinese currency. The standard point of reference in thinking about the role of productivity is the Balassa-Samuelson theory, which focuses on the difference between growth rates of productivity in the traded (e.g., manufacturing) and nontraded sectors (e.g., construction, retail). The Balassa-Samuelson theory predicts that a currency will become stronger when productivity growth in the traded sector exceeds that in the nontraded sector, when compared to the countries that economy trades with.

We have found that each 1 percent increase in Chinese manufacturing productivity over U.S. productivity results in a half percentage point appreciation in the renminbi against the dollar, in real terms. Increases in nontraded sector productivity depreciates the renminbi, in line with the predictions of the Balassa-Samuelson theory. This results holds regardless of how we measure manufacturing productivity.

Depending upon how we measure productivity and upon several other assumptions, we find the renminbi in 2004 to be undervalued 6.1 percent at

the most and 1.4 percent at the least.

These counter-intuitive results suggest that something may be missing from this approach. This framework assumes the relative production costs solely determine the relative prices of tradable and nontradable goods. This condition makes sense when one is talking about “frictionless” economies, wherein workers and capital can move from industry to industry easily. It also requires that over the long term, people’s preferences for certain goods—say manicures versus rice—do not change even as per-capita income rises. Since this is unlikely to be the case, the real exchange rate might change for reasons apart from differing trends in productivity growth.

The key difficulty with this approach is that it relies upon the relationship holding over the sample period. If the entire sample period were one in which the Chinese economy were adjusting to a condition where the Balassa-Samuelson hypothesis held—without actually achieving that condition—then this approach would be invalid. This is not a problem specific to this approach. It also occurs in cases where one is empirically validating purchasing power parity in levels using not actual prices, but price indexes. The

Depending upon how we measure productivity and upon several other assumptions, we find the renminbi in 2004 to be undervalued 6.1 percent at the most and 1.4 percent at the least.

The La Follette School of Public Affairs
at the University of Wisconsin-Madison

Congratulates Faculty Members

Carolyn Heinrich

editor of the *Journal of Public Administration Research and Theory*,
ranked No.1 for citation impact for 2005

Graham Wilson

who completed his term this year as co-editor of *Governance*,
ranked No. 2 for citation impact for 2005

Charles Engel

co-editor of the *Journal of International Economics*

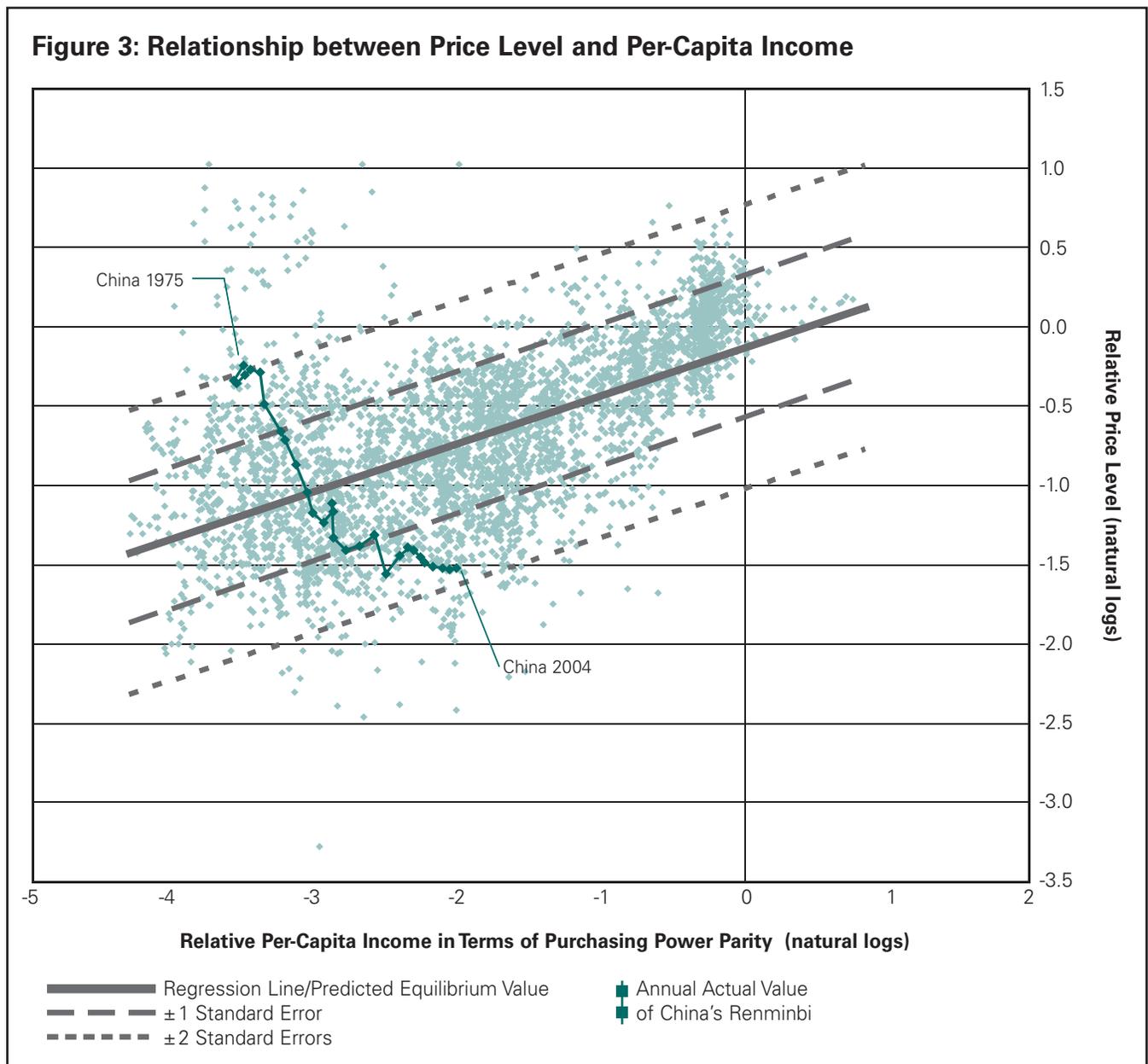
limitation of such approaches, based upon indices, motivates the use of measures where price levels can be directly compared.

A real-world manifestation is highlighted by the well-known “Big Mac” index that *The Economist* publishes. This index measures the dollar price of a Big Mac in different currencies, and, after converting them by the official exchange rate, determines whether the dollar price of a Big Mac in Beijing deviates from that in New York. For instance, on May 22, the price of a Big Mac in Beijing was 10.5 yuan, or \$1.31 after converting by the exchange rate at that date. But a Big Mac in the United States costs \$3.10. That means that one would need to give up 2.37 Big Macs in China to obtain one Big Mac in

America, instead of a one-to-one figure one might expect if it was costless to buy and sell Big Macs across the Pacific. To figure out the undervaluation, consider which exchange rate would set the price of the Big Mac’s equal. It’s 3.39 yuan per dollar. The actual exchange rate was 8.03 yuan per dollar. The difference is 58 percent.

This same type of correlation is observed for price levels associated with bundles of goods—rather than Big Macs—across countries. Why this phenomenon occurs is not clear. One prominent hypothesis is that when per-capita incomes are higher, demand for nontraded goods is higher and this results in a higher price for the total bundle. A higher price for a bundle of goods in the

Figure 3: Relationship between Price Level and Per-Capita Income



United States measured this way means the value of the dollar is higher.

We appeal to this systematic relationship to see whether China is an outlier with respect to this relationship between price levels and per-capita income. If the Chinese price level is substantially below what the relationship predicts, one might say that the Chinese currency is undervalued. In Figure 3 we provide a graphical depiction of the actual (light green dots) vs. predicted (solid gray line) values with prediction intervals. The wide dispersion of observations in the scatterplot should give pause to those who would make strong statements regarding the exact degree of misalignment.

It is interesting to consider the path that the renminbi has traced out in Figure 3. It starts as overvalued at the beginning of the sample, and, over the subsequent three decades, moves toward the predicted equilibrium value and then overshoots, so that by 2004, it is substantially undervalued, by up to 53 percent in level terms.

While the estimated degree of undervaluation is quite substantial, the fact that the end point of the renminbi in 2004 is still close to the conditional mean (the regression line) indicates that one can't be certain that the renminbi is actually undervalued, in statistical terms.

The Interaction of Capital Controls and Corruption

We now augment this simple model with some additional factors. One commonly heard argument is that the Chinese economy is special—namely it is one that is characterized by extreme corruption. Moreover, an extensive capital control regime, including restrictions upon which assets can be purchased in China, and which profits can be repatriated across borders, is still in place. We investigate whether these two aspects of the Chinese economy are of measurable importance in the determination of exchange rates, and if so, whether our conclusions regarding renminbi misalignment are altered as a consequence.

We augment the basic real exchange rate-relative income relationship with the Chinn-Ito capital account openness index, which measures the extent of capital controls. The higher the index, the fewer the impediments to capital flows. In addition we use the *International Country Risk Guide's* Corruption Index, where higher values of the index denote less corruption.

We find that the estimated effect of income on the exchange rate is pretty much the same as we found before. In addition, we find that both capital account openness and the absence of corruption tend to strengthen a currency. These results imply a smaller undervaluation. In other words, to the extent that lack of transparency is a given at any point of time, the renminbi is less undervalued than would be indicated simply from an inspection of relative price levels.

These estimates seem to fly against the view that eliminating Chinese capital controls would cause the renminbi to decline in value. We find in our empirical work that increasing capital account openness when corruption remains high will lead to a weaker currency.

**Given the price of a Big Mac
in Beijing and in New York City,
one would need to give up 2.37
Big Macs in China to obtain
one Big Mac in America.**

Some Additional Thoughts on Measuring Misalignments

The finding that capital account openness, the lack of corruption, and the interaction between the two matter for the level of the exchange rate suggests that our understanding of when a currency is misaligned is highly circumscribed. Experience may be helpful in illustrating this point. It is now widely acknowledged that by many conventional measures—largely based upon relative purchasing power parity—the Thai and Korean currencies did not appear terribly overvalued on the eve of the 1997 crises. For instance, work by Menzie Chinn indicates single-digit overvaluations of the Thai baht and perhaps 10 percent misalignment for the Korean won, based upon relative purchasing power parity measures. Yet, in light of the large contingent liabilities—deposits at banks that fail or corporations that become insolvent due to borrowings in foreign currency—that were subsequently uncovered, the baht and won appeared, in retrospect, quite overvalued. A similar argument can be made with respect to China: According to relative prices, the renminbi may appear undervalued, but, in the context of large stocks of contingent liabilities, the reverse may be true. Key among the candidate contingent liabilities is the large stock of nonperforming loans in the Chinese banking system.

What determines the buildup of contingent liabilities remains unresolved. However, experts generally agree that the lack of transparency (loosely, the

presence of corruption) contributes to the accumulation of such liabilities and hence sets the stage for financial crises. The widely acknowledged lack of transparency in the Chinese political economy appears relevant to thinking about the fair value of that country's currency.

The Policy Context

The preceding discussion has been somewhat academic, in that we focus on measuring the degree of currency misalignment. But policymakers operate in the here and now of the real world. They need to take as given certain conditions, with the understanding that change along certain dimensions is unlikely to occur rapidly. Chief among these circumstances is capital controls and the distortions in the Chinese financial system. Also taken as given is rapid economic growth in China. *Under these conditions*, the burgeoning trade balance and current account surpluses and foreign exchange reserves approaching a trillion U.S. dollars can be taken as indicators of renminbi undervaluation.

It is reasonable, then, for policymakers to ask for some currency adjustment, perhaps even more than what has taken place to date. At the same time, two important worries should be heeded. The first is that renminbi appreciation in and of itself is unlikely to alter the basic problem of a massive and expanding U.S. trade deficit. That problem is first and foremost a “made-in-America” issue driven by collapsed household and public sector savings and by heavy dependence upon imported oil.

Renminbi appreciation in and of itself is unlikely to alter the basic problem of a massive and expanding U.S. trade deficit. That problem is first and foremost a “made-in-America” issue.

The second worry is that conditions can change, and change quickly. Indeed, with the U.S. economy slowing and monetary policy increasingly restrictive in the developed economies,

the possibility of fairly abrupt changes in the emerging economies—including in China—should not be dismissed lightly.

In sum, the renminbi is probably misaligned by at least one or two criteria. But it would be an enormous mistake to think a stronger renminbi is a panacea for what ails America (or China, for that matter). ♦



La Follette
Policy Report

The Robert M. La Follette
School of Public Affairs
University of Wisconsin–Madison
1225 Observatory Drive
Madison, WI 53706

Nonprofit Org.
U.S. Postage
Paid
Madison, WI
Permit No. 658