

Second Quarter 2008

# **Economic** perspectives

- 
- 2** Does education improve health? A reexamination of the evidence from compulsory schooling laws

*Bhashkar Mazumder*

- 
- 17** How do EITC recipients spend their refunds?

*Andrew Goodman-Bacon and Leslie McGranahan*

- 
- 33** Are inflation targets good inflation forecasts?

*Marie Diron and Benoît Mojon*

# Economic perspectives

---

**President**

Charles L. Evans

**Senior Vice President and Director of Research**

Daniel G. Sullivan

**Research Department****Financial Studies**

Douglas Evanoff, Vice President

**Macroeconomic Policy Research**

Jonas Fisher, Economic Advisor and Team Leader

**Microeconomic Policy**

Daniel Aaronson, Economic Advisor and Team Leader

**Payment Studies**

Richard Porter, Vice President

**Regional Programs**

William A. Testa, Vice President

**Economics Editor**

Anna L. Paulson, Senior Financial Economist

**Editor**

Helen O'D. Koshy

**Associate Editors**

Kathryn Moran

Han Y. Choi

**Graphics**

Rita Molloy

**Production**

Julia Baker

**Economic Perspectives** is published by the Research Department of the Federal Reserve Bank of Chicago. The views expressed are the authors' and do not necessarily reflect the views of the Federal Reserve Bank of Chicago or the Federal Reserve System.

© 2008 Federal Reserve Bank of Chicago

**Economic Perspectives** articles may be reproduced in whole or in part, provided the articles are not reproduced or distributed for commercial gain and provided the source is appropriately credited. Prior written permission must be obtained for any other reproduction, distribution, republication, or creation of derivative works of **Economic Perspectives** articles. To request permission, please contact Helen Koshy, senior editor, at 312-322-5830 or email [Helen.Koshy@chi.frb.org](mailto:Helen.Koshy@chi.frb.org).

**Economic Perspectives** and other Bank publications are available on the World Wide Web at [www.chicagofed.org](http://www.chicagofed.org).

 **chicagofed.org**

ISSN 0164-0682

# Contents

---

Second Quarter 2008, Volume XXXII, Issue 2

---

## **2** Does education improve health? A reexamination of the evidence from compulsory schooling laws

**Bhashkar Mazumder**

This article analyzes the impact of compulsory schooling laws early in the twentieth century on long-term health. The author finds no compelling evidence for a causal link between education and health using this research design. Further, the results suggest that only a small fraction of health conditions are affected by education, and several of those are conditions, such as sight and hearing, where economic theories don't appear to be relevant.

---

## **17** How do EITC recipients spend their refunds?

**Andrew Goodman-Bacon and Leslie McGranahan**

The authors determine what items are purchased using the earned income tax credit (EITC)—one of the largest sources of public support for lower-income working families in the U.S. They find that recipient households' EITC payments are used primarily for vehicle purchases and transportation spending, both of which are crucial to job access and consistent with the EITC's prowork goals.

---

## **33** Are inflation targets good inflation forecasts?

**Marie Diron and Benoît Mojon**

The authors show that quantified inflation objectives, which have been adopted by many industrialized countries, can be used as rule-of-thumb forecasting devices. Remarkably, they yield smaller forecast errors than widely used forecasting models and the forecasts of professional experts.

---

## **46** *International Banking Conference The Credit Market Turmoil of 2007–08: Implications for Public Policy*



# Does education improve health? A reexamination of the evidence from compulsory schooling laws

**Bhashkar Mazumder**

## Introduction and summary

Improving the long-term health of the population is clearly an important goal for policymakers. It is also likely to become even more so in the coming years with the aging of the baby boomers and the anticipated health-related costs that will accompany this demographic change. Therefore, understanding which policy levers might improve health is of interest. In a provocatively titled front page article, “A surprising secret to a long life: Stay in school,” the *New York Times* recently suggested that many researchers now believe that education is the key factor in promoting health.<sup>1</sup> While social scientists have long known that there is a strong positive *correlation* between education and longevity, many researchers have speculated that this association was not truly *causal*, meaning one didn’t necessarily lead to the other. Rather, the link was thought to reflect either the fact that for a variety of other reasons (for example, parental income and personal attitudes), people who tend to acquire more schooling also tend to be in better health, or that healthier children stayed in school longer. Of course, in the absence of evidence of a causal link, there is no reason to expect that policies aimed at increasing educational attainment will result in improvements in health.

The *New York Times* article was based upon the results of a recent study by economist Adriana Lleras-Muney (2005) that provides perhaps the strongest evidence to date that education has a causal effect on health. By implementing an instrumental variables (IV) strategy, this research analyzes changes in compulsory schooling and child labor laws across different states early in the twentieth century and uses this information to infer the effects of education on mortality. The idea behind this strategy is that if differences in these laws induced people born in different states in different years to obtain different levels of schooling for reasons that are unrelated to any other

determinants of health, then one can estimate a true causal effect that is not confounded by the other factors. Lleras-Muney finds that increased schooling due to these laws led to dramatic reductions in mortality rates during the 1960s and 1970s. In fact, the results imply that one more year of schooling would lower the mortality rate over a ten-year period by nearly 60 percent—a result that is perhaps implausibly large.

If it is true that more education leads to improved health, such a finding also raises a second important question—namely, *how*, exactly, does education affect health? Economists have proposed a variety of theories including: that more education leads to better jobs and more financial resources; that education improves knowledge and decision-making ability, which improves health; and that education influences other kinds of behavioral responses that, in turn, lead to better health outcomes. So far, however, there is little convincing empirical evidence on how to evaluate the importance of these factors.

In this article, I reexamine the use of these compulsory schooling laws as a way of identifying the causal effects of education on health through the IV approach. Given the fundamental importance of the question of whether more education is causally linked to better health, it is worth investigating the robustness of the relationship. I estimate the same types of models used in the earlier research, using a much larger sample and improved measures of compulsory schooling laws. I also present alternative specifications of the statistical model that may better account for other

*Bhashkar Mazumder is a senior economist in the Economic Research Department and the executive director of the Chicago Research Data Center at the Federal Reserve Bank of Chicago. The author thanks Douglas Almond, Claudia Goldin, Adriana Lleras-Muney, Anna Paulson, and Diane Schanzenbach.*

reforms that were going on during the same period. For example, during the early period of the twentieth century, there were fairly dramatic improvements in public health measures that led to large declines in concurrent mortality (Cutler and Miller, 2005). For school-age children specifically, new nutrition and vaccination programs may have resulted in improved long-term health, independent of any effects of increased education.

In addition, if compulsory schooling laws can be used to identify a causal relationship, then they also ought to be useful in identifying *how* education improves health. This can be analyzed by using data on very specific health conditions for which existing theories might favor one explanation versus another. For example, if processing information and decision-making ability are the critical channels by which education affects health, then we might expect lower incidences of chronic diseases, such as arthritis, cancer, diabetes, lung disease, and heart disease. These are conditions that might respond better to more sophisticated management plans or behavioral changes. If the key factor is increased access to high-quality health care due to greater financial resources, then we might expect that a broad range of health outcomes would be improved. Therefore, it makes sense to apply the same methodology to other outcomes besides mortality.

A careful analysis of how education affects health using the IV approach also serves as a credibility check on the methodology. If, for example, all of the health effects appeared to be related to the long-term effects of poor nutrition, then a plausible alternative hypothesis would be that changes in compulsory schooling laws are really just picking up the long-term health effects of improved nutrition in schools. In that case, the assumption that these laws represent exogenous sources of schooling differences would be invalid, and the estimates would not represent a causal relationship between education and health.

In order to address these issues, I first reexamine the effects of education on mortality from Lleras-Muney (2005, 2006) by replicating the results and extending them by adding significantly more data and employing a variety of robustness checks. I find that the effect of education on mortality is not robust to the inclusion of state-specific time trends, casting doubt on whether there is a true causal effect. At a minimum, my results show that the point estimates are much smaller than those previously found in the literature. Moreover, the results appear to be driven by the earliest cohorts (born in 1901–12) during the 1960–70 period.

Second, I use individual-level data on health outcomes from the U.S. Census Bureau's *Survey of*

*Income and Program Participation* (SIPP) to further investigate the causal pathways between compulsory schooling and health. In contrast to the U.S. Census data, which requires the use of a cohort grouping strategy to infer mortality, the SIPP provides data on the health status of each individual so that we can be sure that those who were affected by the compulsory schooling laws are indeed the same individuals registering the change in health. Using the SIPP with the same IV strategy, I find large and statistically significant effects of education on general health status that are robust to the inclusion of state-specific time trends. This suggests that the SIPP micro data are able to overcome the limitations of the U.S. Census data.

However, when I turn to the results that identify which specific health conditions were affected by education improvements induced by compulsory schooling laws, the results do not point to a coherent story of how education affects health. For example, only a small fraction of health conditions are affected by education, and several of those affected are conditions, such as sight and hearing, where economic theories don't appear to be relevant. What is also striking is the absence of effects among many chronic diseases where decision-making ability is believed pivotal. A limitation of the data, however, is that specific conditions are only identified for a subset of the sample that report having some health limitations. Nevertheless, this pattern of results suggests that the use of compulsory schooling laws as an instrument may be suspect. I also note that in a recent working paper, Clark and Royer (2007) use an even more sophisticated approach to analyze the effects of compulsory schooling law changes in the United Kingdom on mortality. Their findings also cast doubt on whether there is a strong causal connection between education and health.

## Background and previous literature

Kitagawa and Hauser (1973) were the first to document the sharp differences in health in the United States by socioeconomic status. A large number of studies have since replicated this basic finding of a "gradient" in health by education or income, and this pattern has also been found in other countries.<sup>2</sup> For policymakers, a critical question is whether this gradient reflects a causal relationship that can be exploited to improve the long-term health of the population. For example, in a document soliciting research proposals on the pathways linking education to health, the National Institutes of Health (2003) cautioned that: "The association or pathway between formal education and either important health behaviors or diseases may not be causal. Instead it may reflect the

influence of confounding or co-existing determinants or may be bi-directional.”

A review of the literature on whether the education gradient in health is causal may be found in Grossman (2005). While these studies typically find an effect of more education leading to better health, in most cases it is questionable whether the instruments are truly exogenous. For example, Dhir and Leigh (1997) use parent schooling, parent income, and state of residence as instruments, all of which could plausibly affect long-term health independently of their effects through schooling. The innovation by Lleras-Muney (2005) to use changes in compulsory schooling laws early in the twentieth century appears to be more compelling, since it is more plausibly exogenous than instruments used in prior work. Nevertheless, other changes in public policy that coincided with changes in compulsory schooling laws might have led to long-run improvements in health. Cutler and Miller (2005) find that the introduction of clean water technologies during this period could explain as much as half of the concurrent decline in mortality. Similarly, many states introduced food programs in schools, recognizing that compulsory schooling was pointless if children were malnourished. Near the beginning of the twentieth century, Robert Hunter (1904) wrote in the book *Poverty*: “There must be thousands—very likely sixty or seventy thousand children—in New York City alone who often arrive at school hungry and unfitted to do well the work required. It is utterly folly, from the point of view of learning, to have a compulsory school law which compels children, in that weak physical and mental state which results from poverty, to drag themselves to school and to sit at their desks, day in and day out, for several years, learning little or nothing.” In response to this situation, Philadelphia, Boston, Milwaukee, New York, Cleveland, Cincinnati, and St. Louis all began large-scale programs to provide food in public schools during the 1900s and 1910s (Gunderson, 1971). Mazumder (2007) also provides suggestive evidence that the mechanism by which compulsory schooling laws might have improved long-term health was through school requirements for vaccination against smallpox. If improvements in nutrition and vaccination programs were coincident with changes in compulsory schooling laws, then these might explain some or all of the long-term health improvements that were associated with changes in these laws.

Supposing that it is true that more education leads to improved health, this finding raises an interesting question—namely, *how*, exactly, does education affect health? As Richard Suzman of the National Institute on Aging recently stated, “Education ... is a particularly powerful

factor in both life expectancy and health expectancy, though truthfully, we’re not quite sure why.”<sup>23</sup> Economists have proposed a variety of explanations. These theories typically emphasize the role of education in affecting various proximate determinants of health, including financial resources, knowledge and decision-making ability, and other behavioral characteristics that could lead to better health outcomes.

Financial resources come into play because better educated individuals may obtain higher paying and more stable jobs and thereby may be able to afford better quality health care and health insurance. With greater economic resources, they may also choose safer and more secure living and work environments. One might expect that if financial resources are the key factor behind the link between education and health, then we should expect to see virtually all forms of health conditions affected by exogenous sources of increased education.

The second explanation is that higher levels of schooling may lead to greater knowledge and an improved ability to process information and make better choices or take better advantage of technological improvements. In one widely cited paper, Goldman and Smith (2002) note that better educated patients may manage chronic conditions better. Those with more schooling adhere more closely to treatment regimens for human immunodeficiency virus (HIV) infection and diabetes, which can be fairly complex. For such conditions, the ability to form independent judgments and comprehend treatments is important, and apparently is fostered by schooling. Accordingly, Goldman and Smith (2002, p. 10934) state that “self-maintenance is an important reason for the very steep SES [socioeconomic status] gradient in health outcomes.” Glied and Lleras-Muney (2003) argue that “the most educated make the best initial use of new information about different aspects of health,” permitting them to respond more adeptly to evolving medical technologies.

Finally, it could be that education induces other kinds of behavioral changes. For example, the better educated may value the future more than the present compared with those with less education, and therefore, the better educated may take better care of their health (Becker and Mulligan, 1997). Others have argued that education improves one’s perception of one’s *relative* status in society and that improved social standing is associated with better health (Marmot, 1994).

### **Mortality analysis: Methodology and data**

The first part of the analysis estimates the effects of education on mortality, using the approach developed by Lleras-Muney (2005). In the absence of a

large sample of data on individuals containing both education and lifespan, I use group-level data from successive U.S. *Decennial Censuses* to estimate mortality rates. Specifically, I use population estimates for groups defined by state of birth, gender, and year of birth to estimate the mortality rate across ten-year periods. The mortality rate at time  $t$  for birth cohort  $c$  of gender  $g$  born in state  $s$ , ( $M_{cgst}$ ), is simply measured as the percentage decline in the population count ( $N_{cgst}$ ) within these cells over the subsequent ten years:

$$1) \quad M_{cgst} = \frac{N_{cgst} - N_{cgst+1}}{N_{cgst}}.$$

I then model the mortality rate for each cell as follows:

$$2) \quad M_{cgst} = a + E_{cgst} \pi + W_{cs} \delta + \gamma_c + \alpha_s + \theta_{cr} + fem + \tau_t + \varepsilon_{cgst},$$

where  $E_{cgst}$  is the average education level for that cell at time  $t$  and  $W_{cs}$  measures a set of cohort and state-specific controls measured at age 14 intended to capture differences in other potential early life determinants of mortality (for example, manufacturing share of employment and doctors per capita). The model also includes a set of cohort dummies  $c$ , state of birth dummies  $s$ , interactions between cohort and region of birth  $\theta_{cr}$ , a female dummy ( $fem$ ), and year dummies  $\tau_t$ .

One straightforward way to estimate  $\pi$  in equation 2 would be through weighted least squares (WLS), with the weights corresponding to the population represented by each cell. However, this would produce a biased estimate because of omitted variables. Any number of factors could plausibly be associated with both higher education and lower mortality even at the group level. Therefore, I use two-stage least squares, where in the first stage, education is instrumented with the set of compulsory schooling laws,  $CL_{cs}$ , in place for each cohort and state of birth:

$$3) \quad E_{cgst} = b + CL_{cs} \rho + X_{cgst} \beta + W_{cs} \delta + \gamma_c + \alpha_s + \theta_{cr} + fem + \tau_t + u_{cgst}.$$

In Lleras-Muney (2005), the instruments for the compulsory schooling laws were constructed in the following way. The variable *childcom* measured the minimum required age for work minus the maximum age before a child is required to enter school, by state of birth and by the year the cohort is age 14. This

variable takes on one of eight values. A set of indicator variables were then used as instruments. In addition, an indicator for whether school continuation laws were in place in that state was also used. These laws required workers of school age to continue school part time. However, it probably makes more sense to match individuals to the laws concerning the maximum age for school *entry* around the age at which students start school, rather than to the laws in place when they were age 14. Therefore, I use a different set of data independently collected by Goldin and Katz (2003).<sup>4</sup> Goldin and Katz carefully compared their series with other codings of the compulsory schooling laws (for example, Lleras-Muney, 2005; and Acemoglu and Angrist, 2001) and resolved differences wherever possible. Since the Goldin and Katz data go back further in time, it is possible to match all of the cohorts to the school entry age laws in effect when the cohorts were younger than 14. I use these data to measure the required age for school entry when the cohorts were at age 8 instead of 14. In principle, incorporating these data should provide a better measure of the total years of compulsory schooling.

Several estimation samples are constructed for this part of the analysis. Initially, I produce a sample combining data from the 1 percent Integrated Public Use Microdata Series (IPUMS) from the 1960, 1970, and 1980 U.S. Censuses in order to replicate the basic results in Lleras-Muney (2005, 2006).<sup>5</sup> I then expand the analysis in stages. First, I replace the 1 percent samples in 1970 and 1980 with a 2 percent sample for 1970 and a 5 percent sample for 1980. Second, I also expand the periods by adding 5 percent samples for 1990 and 2000. Following the literature, I restrict the analysis to cohorts born between 1901 and 1925, top-code years of education at 18 starting in 1980, and exclude immigrants and blacks.<sup>6</sup> For the expanded samples, I also exclude cases where age, state of birth, and education are imputed by the U.S. Census Bureau. The descriptive statistics for the replication sample and the expanded sample are shown in table 1.

It is worth noting that the death rate for the 1970–80 period is quite a bit larger with the expanded sample but that the standard deviation is about 20 percent lower. There are now also five additional cells that had missing data when using just the 1 percent samples. The death rates for the 1980–90 and 1990–2000 periods are much higher because I follow these same cohorts when they are much older. Figure 1 plots the death rates by age for each U.S. Census year. This highlights the importance of controlling for age in the specifications, which is done by adding polynomials in age to the models.

**TABLE 1**

**Summary statistics for Integrated Public Use Microdata Series samples**

Variables	1960 1%, 1970 1%, and 1980 1% samples			1960 1%, 1970 2%, 1980 5%, 1990 5%, and 2000 5% samples		
	Mean	Standard deviation	Number of observations	Mean	Standard deviation	Number of observations
<b>Ten year death rates</b>						
Overall	0.108	0.136	4,792	0.213	0.173	8,636
1960–70	0.110	0.119	2,395	0.113	0.105	2,397
1970–80	0.105	0.152	2,397	0.154	0.125	2,400
1980–90	—	—	—	0.287	0.170	2,399
1990–2000	—	—	—	0.433	0.122	1,440
<b>Individual characteristics</b>						
Education	10.548	0.990	4,795	10.729	1.002	8,636
1960 dummy	0.471	0.499	4,795	0.325	0.469	8,636
1970 dummy	—	—	—	0.289	0.453	8,636
1990 dummy	—	—	—	0.142	0.349	8,636
Female	0.517	0.500	4,795	0.532	0.499	8,636
Age	50.366	8.482	4,795	56.811	11.287	8,636
Born in 1905	0.031	0.174	4,795	0.025	0.157	8,636
Born in 1910	0.038	0.191	4,795	0.031	0.174	8,636
Born in 1915	0.044	0.205	4,795	0.047	0.211	8,636
Born in 1920	0.048	0.213	4,795	0.052	0.222	8,636
Born in 1925	0.050	0.217	4,795	0.057	0.232	8,636
<b>State of birth characteristics</b>						
Percentage urban	53.523	21.279	4,795	53.778	21.153	8,636
Percentage foreign-born	11.737	8.523	4,795	11.562	8.430	8,636
Percentage black	8.983	11.901	4,795	8.945	11.787	8,636
Percentage employed in manufacturing	0.067	0.038	4,795	0.066	0.037	8,636
Annual manufacturing wage (\$)	7,171.39	1,343.09	4,795	7,206.15	1,353.57	8,636
Value of farm per acre (\$)	540.05	276.35	4,795	535.18	272.57	8,636
Per capita number of doctors	0.001	0.000	4,795	0.001	0.000	8,636
Per capita education expenditures (\$)	97.01	42.05	4,795	99.78	41.71	8,636
Number of school buildings per square mile	0.174	0.090	4,795	0.172	0.090	8,636

Notes: Summary statistics are for state of birth, cohort, and gender cells. All means and standard deviations use sample weights where the weights are the population estimates for the cell in the base period.

Source: Author's calculations based on data from the University of Minnesota, Minnesota Population Center, Integrated Public Use Microdata Series.

**Health analysis: Methodology and data**

The methodological approach changes only slightly when I turn to using individual-level data from the SIPP. Many of the outcomes in the SIPP are indicator variables that take on the value of 1 if a particular health problem is present and 0 otherwise. Therefore, I now use two-stage conditional maximum likelihood, or 2SCML (Rivers and Vuong, 1988), rather than IV.<sup>7</sup> Rivers and Vuong show that 2SCML has desirable statistical properties, is easy to implement, and produces a simple test for exogeneity. I continue to use IV for the few continuous dependent variables. Also, all of the analysis is now done using individual-level data. The statistical model is similar to equation 2, only now I use the latent variable framework:

$$4) \quad y_{it}^* = a + E_i\pi + X_i\beta + W_{cs}\delta + \gamma_c + \alpha_s + trend_s + \tau_t + fem + \varepsilon_{it}$$

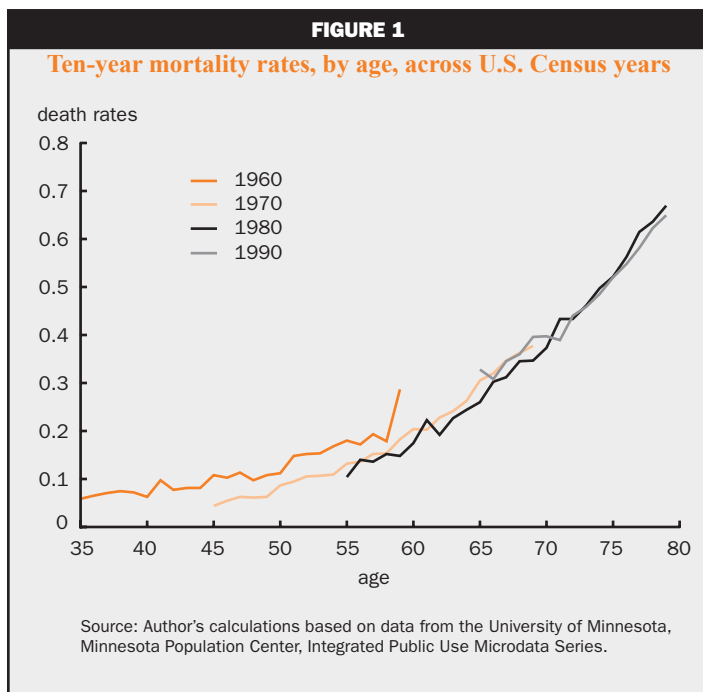
$$5) \quad y_{it} = 1 \text{ if } y_{it}^* > 0, \quad y_{it} = 0 \text{ if } y_{it}^* \leq 0.$$

In the first stage, I run a similar regression as before:

$$6) \quad E_i = b + CL_{cs}\rho + X_i\beta + W_{cs}\delta + \gamma_c + \alpha_s + trend_s + \tau_t + d + \varepsilon_{it}$$

To implement 2SCML, I use the predicted residuals from equation 6,  $\hat{\varepsilon}_{it}$ , and I include it as an additional right-hand side variable (along with the actual value of  $E_i$ ) when running the second stage probit. For comparability, I use the same sample restrictions and





covariates as in the U.S. Census results, with only a few exceptions. I include a quadratic in age and use state-specific cohort trends to address concerns that region of birth interacted with cohort may not adequately control for state-specific factors that are smoothly changing over time.<sup>8</sup>

The sample is constructed by pooling individuals from the 1984, 1986–88, 1990–93, and 1996 SIPP panels. Each SIPP panel surveys approximately 20,000 to 40,000 households, and most panels are representative of the noninstitutionalized population.<sup>9</sup> Because participation in many programs is closely related to an individual's health and disability status, the SIPP routinely collects information on health and medical conditions. The SIPP is also ideally suited for this analysis because it contains the state of birth of all sample members, which allows me to implement the IV strategy of using compulsory schooling laws during childhood.

One especially useful outcome is self-reported health (SRH). The SRH is on a 1–5 scale, where 1 is “excellent,” 2 is “very good,” 3 is “good,” 4 is “fair,” and 5 is “poor.” The SRH has been found to be an excellent predictor of mortality and changes in functional abilities among the elderly (Case, Lubotsky, and Paxson, 2002). I experiment with this measure in a few ways. First, I use it as a continuous variable. Second, I use indicators for being in poor health or in fair or poor health. Finally, I use the health utility scale that measures the differences between the categories in a health model using the *National Health Interview Survey*

(conducted by the U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics).<sup>10</sup>

I also examine some other general outcomes. These are whether the individual was hospitalized during the past year, the number of times she was hospitalized, the total number of nights spent in the hospital, and the number of days spent in bed in the past four months.

There are also questions dealing with functional activities, activities of daily living, and instrumental activities of daily living that are derived from the International Classification of Impairments, Disabilities, and Handicaps (ICIDH). I assembled a common set of questions that were consistently asked across surveys. These are whether the individual has “difficulty” with seeing, hearing, speaking, lifting, walking, and climbing stairs, as well as whether the person can perform any of these activities “at all.” In addition, there is information on whether individuals have difficulty getting around inside the house, going outside of the house, or getting in or out of bed, as well as whether they need the assistance of others for these activities.

For a subset of individuals who report limited abilities in certain tasks or who have been classified as having a work disability (“health limitation”), detailed information is collected on a number of very specific health conditions including: arthritis or rheumatism; back or spine problems; blindness or vision problems; cancer; deafness or serious trouble hearing; diabetes; heart trouble; hernia; high blood pressure (hypertension); kidney stones or chronic kidney trouble; mental illness; missing limbs; lung problems; paralysis; senility/dementia/Alzheimer's disease; stiffness or deformity of limbs; stomach trouble; stroke; thyroid trouble or goiter; tumors (cyst or growth); or other.<sup>11</sup> Since the specific health ailments are only asked of specific subsamples, they probably only pick up on the most severe cases. Even though many of the sample individuals are not actually asked about these specific health conditions, I still include them in the estimation sample so that the sample is not a selected sample of only those in poor health. The summary statistics for these data are shown in table 2.

**Mortality results**

I begin by trying to match the estimates of the effect of education on ten-year mortality rates shown in

TABLE 2

## Summary statistics for Survey of Income and Program Participation sample

Variables	Mean	Standard deviation	Number of observations
<b>Outcomes</b>			
Self-reported health (1 is excellent, 5 is poor)	3.084	1.138	26,030
Poor health	0.119	0.324	26,030
Fair or poor health	0.357	0.479	26,030
Health index (1–100 scale)	67.992	24.842	26,030
Hospitalized in last year	0.180	0.384	26,484
Days in bed, last four months	3.937	17.030	25,223
Number of times hospitalized	0.282	1.029	22,229
Number of nights in hospital	1.908	7.898	26,274
Trouble seeing	0.136	0.342	20,853
Trouble hearing	0.152	0.359	20,845
Trouble speaking	0.021	0.144	20,834
Trouble lifting	0.237	0.425	20,837
Trouble walking	0.289	0.453	20,799
Trouble with stairs	0.276	0.447	20,820
Trouble getting around outside the home	0.129	0.335	17,401
Trouble getting around inside the home	0.059	0.235	17,643
Trouble getting in/out of bed	0.079	0.270	17,636
Trouble seeing at all	0.023	0.149	20,811
Trouble hearing at all	0.013	0.114	20,819
Trouble speaking at all	0.003	0.052	15,138
Trouble lifting at all	0.115	0.319	20,789
Trouble walking at all	0.154	0.361	20,723
Trouble with stairs at all	0.116	0.321	20,775
Needs help getting around outside	0.088	0.283	13,610
Needs help getting around inside	0.024	0.154	13,893
Needs help getting in/out of bed	0.025	0.156	13,868
Work limitation due to health conditions	0.423	0.494	19,073
Arthritis	0.129	0.335	19,073
Back	0.062	0.242	19,073
Blind	0.026	0.159	19,073
Cancer	0.016	0.125	19,073
Deaf	0.023	0.149	19,073
Deformity	0.027	0.162	19,073
Diabetes	0.030	0.170	19,073
Heart	0.090	0.287	19,073
Hernia	0.006	0.080	19,073
Hypertension	0.036	0.185	19,073
Kidney	0.005	0.067	19,073
Lung	0.043	0.203	19,073
Mental illness	0.005	0.067	19,073
Missing limb	0.003	0.056	19,073
Paralysis	0.006	0.075	19,073
Senility	0.007	0.084	19,073
Stomach	0.010	0.099	19,073
Stroke	0.021	0.144	19,073
Thyroid	0.003	0.056	19,073
Other	0.066	0.247	19,073
<b>Individual characteristics</b>			
Education	11.432	3.208	26,030
Female	0.580	0.494	4,795
Age	72.079	5.606	4,795

Source: Author's calculations based on data from the U.S. Census Bureau, *Survey of Income and Program Participation*.

Lleras-Muney (2006).<sup>12</sup> Using WLS, Lleras-Muney's estimate is  $-0.036$ , and using IV, her estimate is  $-0.063$ . These estimates imply huge effects. For example, the IV estimate implies that one additional year of education

would reduce the ten-year mortality rate by about 60 percent.<sup>13</sup> In table 3, I show the results of the replication exercise, as well as the effects of expanding the sample and employing additional robustness checks.

TABLE 3

## New estimates of effects of education on mortality

Sample and specification	WLS	IV	Number of observations
<b>A. 1960–80</b>			
1960–1980 1%:			
No age controls, region × cohort	−0.036 (0.004)	−0.072 (0.025)	4,792
1960 1%, 1970 2%, and 1980 5%:			
No age controls, region × cohort	−0.045 (0.004)	−0.045 (0.024)	4,797
With age cubic, region × cohort	−0.039 (0.004)	−0.047 (0.024)	4,797
With age cubic × Census year, region × cohort	−0.040 (0.004)	−0.047 (0.024)	4,797
With age cubic × Census year, state × cohort trend	−0.048 (0.004)	−0.016 (0.024)	4,797
<b>B. 1960–2000</b>			
1960 1%, 1970 2%, and 1980–2000 5%:			
With age cubic × Census year	−0.034 (0.003)	−0.026 (0.015)	8,636
With age cubic × Census year, state × cohort trend	−0.036 (0.003)	−0.012 (0.016)	8,636
<b>C. 1960–2000, by Census year</b>			
1960 1%, 1970 2%, and 1980–2000 5% with age cubic × Census year:			
Estimated effect for 1960–70	−0.025 (0.006)	−0.081 (0.052)	2,397
Estimated effect for 1970–80	−0.061 (0.005)	−0.023 (0.033)	2,400
Estimated effect for 1980–90	−0.043 (0.004)	0.023 (0.029)	2,399
Estimated effect for 1990–2000	−0.012 (0.005)	0.027 (0.039)	1,440
<b>D. 1960–2000, by age</b>			
1960 1%, 1970 2%, and 1980–2000 5% with age cubic × Census year:			
35–54 year olds	−0.017 (0.005)	−0.067 (0.036)	2,879
55–64 year olds	−0.039 (0.005)	0.063 (0.053)	2,398
65–89 year olds	−0.030 (0.003)	−0.047 (0.023)	3,071
<b>E. 1960–2000, by cohort</b>			
1960 1%, 1970 2%, and 1980–2000 5% with age cubic × Census year:			
Cohorts born in 1901–12	−0.019 (0.004)	−0.203 (0.125)	3,644
Cohorts born in 1913–25	−0.017 (0.004)	0.025 (0.023)	4,992

Notes: WLS means weighted least squares. IV means instrumental variables. The dependent variable is the ten-year mortality rate; table entries are the coefficient on education. All specifications include year dummies, cohort dummies, state of birth dummies, region of birth interacted with cohort, and an intercept (except for panel A, fifth row, and panel B, second row). Estimates are weighted using the number of observations in the cell in the base year. Standard errors, shown in parentheses, are clustered at the state of birth and cohort level.

In the first row of panel A of table 3, I match the WLS estimate of  $-0.036$  exactly, although my IV estimate of  $-0.072$  is slightly larger. It is also worth pointing out that the partial  $F$  statistic on the first stage regression is reasonable at 7.5.<sup>14</sup> The second row of panel A uses the 1960 (1 percent) sample, as well as the larger samples for 1970 (2 percent) and 1980 (5 percent), and utilizes the Goldin and Katz (2003) data for constructing the instruments. I find that the WLS estimate rises to  $-0.045$  and that the IV estimates drop considerably to  $-0.045$ . Had I used the Lleras-Muney data for constructing the instruments, the estimate would be exactly the same at  $-0.045$ . However, the standard error would have declined by about 25 percent relative to the first row, suggesting that expanding the sample provides considerably more precision. In the third and fourth rows of panel A, I control for age and find that this lowers the WLS estimates a little and increases the IV estimates a little. In the fifth row, I drop the region of birth interactions with cohort and instead use state-specific linear (cohort) trends. This raises the WLS estimate to  $-0.048$ , but I now find that the IV coefficient is sharply lower at  $-0.016$  and is no longer statistically significant. However, the fact that the standard error does not rise suggests that the precision is the same when including the state-specific trends.

In panel B of table 3, I add data from the 5 percent samples of the 1990 and 2000 U.S. Censuses. With this larger data set, I construct death rates over four ten-year periods and therefore follow cohorts over a longer period with a considerably larger sample. Given that the sample also tracks the cohorts later in life when mortality rates are much higher, the age controls are essential. I use a cubic in age, although I find that the results are not very sensitive to the choice of the polynomial. Since medical technology and other health-related factors might change over time, I have also interacted the cubic in age with the U.S. Census year. In this specification (the first row of panel B), I now find that the WLS estimate is about  $-0.034$  and that the IV estimate is  $-0.026$ . Both of these estimates are a bit more plausible than the ones mentioned previously. The IV estimate is now significant at the 10 percent level, but not at the 5 percent level. With this larger sample, the inclusion of state-specific cohort trends again results in a point estimate that is much smaller in magnitude ( $-0.012$ ) and not statistically distinguishable from zero (the second row of panel B), despite a similar degree of precision.

In the remaining panels of table 3, I examine how the effects vary by year, age, and cohort. In panel C, I separately estimate the education coefficient for each U.S. Census year. Since the specification includes a

full set of cohort dummies, these are equivalent to age controls when using a single U.S. Census year. Although the WLS estimates are significant in all years, they peak in 1970–80 at  $-0.061$  and drop to only  $-0.012$  by 1990–2000. The IV estimates have large standard errors, so they are likely to be imprecisely estimated. Nonetheless, the point estimate is large only for 1960–70 and is actually positive for 1980–90 and 1990–2000. In panel D, I stratify the sample by three age ranges: 35–54, 55–64, and 65–89. Here I observe different patterns between the WLS and IV specifications. The WLS estimates suggest that the largest effect may be for those aged 55–64, while the IV estimates are largest for those aged 35–54. Given the imprecision of the estimates, I cannot draw any meaningful inferences regarding the age pattern.

Panel E of table 3, however, provides a striking result when using the IV specification. It appears that the entire effect of education on mortality arising from compulsory schooling laws is due to cohorts born in 1901–12, who constitute just over 40 percent of the sample. In fact for those born in 1913–25, the point estimate is actually positive.

### Interpreting the mortality results

I interpret the results in the fifth row of panel A and the second row of panel B of table 3 as suggesting that I cannot reject the null hypothesis that the effect of education on mortality is zero. In other words, education has no causal effect on mortality once I adequately control for state time trends. An alternative view might be that once one includes state time trends, the coefficient is smaller but still negative, and that the standard errors are simply too large to estimate the effect precisely, and therefore, I cannot rule out a causal effect. One might be concerned, for example, that the instruments are highly collinear with the time trends. However, as I have shown, the standard errors do not rise when including the time trends. In any case, this alternative interpretation of the results would implicitly start with the hypothesis that there *is* a causal effect and that the results here do not offer sufficient evidence to reject *that* hypothesis—a strong assumption given that the literature has yet to successfully identify a causal effect.

If one takes seriously the point estimates shown in the fifth row of panel A and the second row of panel B of table 3 (despite their statistical insignificance), then this implies that the causal effects of education on mortality are much smaller than previously thought. A more reasonable estimate then is that an additional year of schooling lowers mortality risk over a ten-year period by about 10 percent. This is still a large



effect that might reflect the true causal effect. Still, it bears repeating that using the current research design, I am unable to reject the hypothesis that the true effect is actually zero.

My analysis also suggests that, upon closer inspection, the results are driven by cohorts born very early in the century and their mortality experience during the 1960–70 period. One possible explanation could be that the effect of education stayed roughly constant but that compulsory schooling laws had their biggest effect on those born earlier in the century. However, I have run the first-stage regressions by these cohort groupings and found that the partial  $F$  statistics on the instruments are actually much higher for the 1913–25 cohorts. This suggests that the schooling laws may actually have been more binding for the later cohorts, casting doubt on this alternative explanation.

### Health outcome results

Table 4 presents the results using the microdata on health outcomes using the SIPP. The first column shows the effects of education using a simple probit (or ordinary least squares, or OLS), which does not account for endogeneity. The second column presents the 2SCML (or IV) estimates using the compulsory schooling laws as instruments. Given the possible effects of education on mortality and the fact that outcomes in the SIPP are not observed until at least 1984, one might not expect any remaining health effects to be apparent. As it turns out, I do find significant effects using the instruments for several broad health outcomes. The first row of panel A shows that self-reported health measured as a continuous variable is affected by education. The IV estimate of  $-0.23$  is more than twice the OLS estimate of  $-0.09$ . In the fourth column using a Hausman test of exogeneity, I can reject that the OLS and IV coefficients are the same at the 7 percent level (shown as  $0.074$  in the table). Translating the SRH into a health index on a 1–100 scale following Johnson and Schoeni’s (2007) approach, the IV estimate implies that an increase in schooling by one year improves the health index by 4.5 points, or about 7 percent evaluated at the mean (third column). I also estimate that the probability of being in fair or poor health is reduced by 8.2 percentage points with an additional year of schooling, a fairly large effect that is statistically different from the naive probit at the 18 percent significance level. I do not find, however, that any of the measures of hospitalization or days spent in bed are significant when accounting for endogeneity.

Looking across a variety of measures of physical function, I find that, while all of the naive probit estimates are significant and of the expected sign, the

two-stage estimates are typically not significant. Those who have an additional year of schooling because of compulsory schooling laws are no less likely to have trouble lifting, walking, climbing stairs, getting around outside the house, getting around inside the house, or getting into or out of bed. In fact for many of these outcomes, the coefficients are actually positive, suggesting they have a greater propensity for worse health. On the other hand, those with greater schooling associated with compulsory schooling laws are dramatically less likely to experience problems with seeing, hearing, or speaking. In almost all of these cases, the differences between the simple probit and the 2SCML estimates are very large and statistically different at about the 10 percent level. For example, the 2SCML estimates imply that an additional year of schooling reduces the probability of having trouble “seeing” by 5.6 percentage points. In this sample, the mean rate of this health outcome is 13.6 percent. These results might suggest that the channel by which general health is compromised for those with less schooling may be related to sensory functions.

Next, I estimate results based on the incidence of specific health conditions. Recall that these conditions are only identified for subsets of individuals and that the screening criteria changed across SIPP survey years. Also recall that all individuals are included regardless of whether they were screened for this question, so as to avoid using a sample of only those in poor health. Generally, the underlying health conditions were only asked of individuals who reported particular kinds of activity limitations, reported having a work disability, or reported being in fair or poor health. This is captured by the variable “health limitation,” which, not surprisingly, is significant under both probit and 2SCML. When I turn to the estimated likelihood of having one of the underlying health conditions, the probit estimates once again are significant in every case. The 2SCML estimates, however, are only negative and significant for four outcomes: back or spine problems; stiffness or deformity of a limb; diabetes; and senility/dementia/Alzheimer’s disease. It is important to point out that “trouble seeing,” “trouble hearing,” and “trouble speaking” were never used as screening criteria for asking about an underlying health condition. This likely explains why blindness and deafness are not significant within the subsamples.

Surprisingly, both kidney problems and hypertension appear to be positively associated with more schooling. This is especially notable because these are two outcomes for which self-management and recent technological advances appear to be especially important. According to appendix table B of

TABLE 4

## Estimates of effects of education on health outcomes

Dependent variable	OLS/probit	IV/2SCML	IV/2SCML effect size	Exogeneity test p value	Number of observations
<b>A. General health outcomes</b>					
Self-reported health (1 is excellent, 5 is poor)	-0.0941 (0.0023)	-0.2289 (0.0745)	-0.074	0.074	26,030
Health index (1-100 scale)	1.9674 (0.0511)	4.5345 (1.6738)	0.067	0.131	26,030
Fair or poor health	-0.0359 (0.0010)	-0.0824 (0.0343)	-0.230	0.176	26,030
Poor health	-0.0141 (0.0006)	-0.0269 (0.0206)	-0.226	0.533	26,030
Hospitalized in last year	-0.0049 (0.0008)	-0.0268 (0.0241)	-0.149	0.364	26,484
Days in bed, last four months	-0.3310 (0.0364)	2.1526 (1.4848)	0.547	0.074	25,223
Number of times hospitalized	-0.0101 (0.0024)	-0.0944 (0.0884)	-0.335	0.329	22,229
Number of nights in hospital	-0.0730 (0.0186)	-1.0828 (0.7668)	-0.567	0.185	26,289
<b>B. Functional limitations/activities of daily living/instrumental activities of daily living</b>					
Trouble seeing	-0.0122 (0.0007)	-0.0559 (0.0254)	-0.412	0.085	20,853
Trouble hearing	-0.0103 (0.0007)	-0.0499 (0.0247)	-0.329	0.109	20,845
Trouble speaking	-0.0019 (0.0002)	-0.0192 (0.0079)	-0.909	0.039	20,573
Trouble lifting	-0.0198 (0.0009)	-0.0055 (0.0330)	-0.023	0.667	20,837
Trouble walking	-0.0251 (0.0011)	0.0130 (0.0325)	0.045	0.242	20,797
Trouble with stairs	-0.0250 (0.0010)	-0.0066 (0.0324)	-0.024	0.993	20,820
Trouble getting around outside the home	-0.0120 (0.0008)	-0.0146 (0.0257)	-0.114	0.918	17,401
Trouble getting around inside the home	-0.0048 (0.0005)	0.0051 (0.0208)	0.087	0.635	17,463
Trouble getting in/ out of bed	-0.0056 (0.0006)	0.0013 (0.0230)	0.016	0.764	17,621
Trouble seeing at all	-0.0020 (0.0002)	-0.0078 (0.0084)	-0.343	0.490	20,589
Trouble hearing at all	-0.0008 (0.0001)	-0.0100 (0.0045)	-0.758	0.060	20,256
Trouble speaking at all	0.0000 (0.0001)	-0.0008 (0.0001)	-0.284	0.000	7,516
Trouble lifting at all	-0.0100 (0.0007)	-0.0029 (0.0250)	-0.025	0.775	20,789
Trouble walking at all	-0.0148 (0.0008)	0.0107 (0.0260)	0.069	0.328	20,723
Trouble with stairs at all	-0.0114 (0.0006)	0.0071 (0.0202)	0.061	0.359	20,775
Needs help getting around outside	-0.0066 (0.0007)	0.0044 (0.0153)	0.050	0.470	13,598

TABLE 4 (CONTINUED)

## Estimates of effects of education on health outcomes

Dependent variable	OLS/probit	IV/2SCML	IV/2SCML effect size	Exogeneity test p value	Number of observations
Needs help getting around inside	-0.0010 (0.0002)	0.0108 (0.0078)	0.446	0.125	13,757
Needs help getting in/out of bed	-0.0011 (0.0003)	0.0092 (0.0080)	0.372	0.191	13,794
<b>C. Specific health conditions</b>					
Health limitation	-0.0250 (0.0013)	-0.0743 (0.0348)	-0.175	0.157	19,073
Arthritis	-0.0088 (0.0008)	-0.0043 (0.0217)	-0.034	0.836	19,012
Back	-0.0028 (0.0005)	-0.0349 (0.0167)	-0.561	0.061	18,924
Blind	-0.0014 (0.0003)	0.0145 (0.0084)	0.557	0.060	18,454
Cancer	-0.0007 (0.0002)	0.0025 (0.0078)	0.161	0.677	18,569
Deaf	-0.0003 (0.0002)	-0.0041 (0.0064)	-0.179	0.568	18,422
Deformity	-0.0006 (0.0002)	-0.0159 (0.0066)	-0.591	0.018	18,821
Diabetes	-0.0023 (0.0003)	-0.0258 (0.0082)	-0.868	0.007	18,688
Heart	-0.0062 (0.0006)	-0.0014 (0.0194)	-0.016	0.804	19,025
Hernia	-0.0003 (0.0001)	0.0023 (0.0037)	0.362	0.454	17,179
Hypertension	-0.0031 (0.0004)	0.0376 (0.0124)	1.053	0.000	18,683
Kidney	-0.0001 (0.0001)	0.0042 (0.0027)	0.938	0.072	16,593
Lung	-0.0037 (0.0005)	0.0203 (0.0152)	0.472	0.106	19,060
Mental illness	-0.00009 (0.00008)	-0.0002 (0.0424)	-0.045	0.932	15,794
Missing limb	-0.00007 (0.00005)	-0.0019 (0.0016)	-0.580	0.155	14,565
Paralysis	-0.00011 (0.00006)	0.0016 (0.0020)	0.287	0.348	17,301
Senility	-0.00005 (0.00002)	-0.0015 (0.0006)	-0.214	0.070	17,993
Stomach	-0.0006 (0.0002)	0.0069 (0.0060)	0.695	0.195	17,701
Stroke	-0.0008 (0.0003)	0.0084 (0.0090)	0.397	0.295	18,918
Thyroid	-0.0000001 (0.000000)	0.000001 (0.000000)	0.000	0.000	14,559
Other	-0.0023 (0.0005)	-0.0013 (0.0152)	-0.019	0.947	19,060

Notes: OLS means ordinary least squares. IV means instrumental variables. 2SCML means two-stage conditional maximum likelihood. Standard errors, shown in parentheses, are clustered at the state of birth and cohort level.

Glied and Lleras-Muney (2003), treatment of kidney infections experienced substantial innovation. Among the 56 causes of death, kidney disease experienced the fastest decline in age-adjusted mortality from 1986 to 1995—falling more than 9 percent per year (Glied and Lleras-Muney, 2003, p. 8, appendix table B). Accordingly, a steep (negative) gradient between education and kidney disease would presumably be expected. It is therefore of note that the 2SCML specification finds an increase in the incidence of kidney problems among those with high education. Treatment of diabetes is “often considered the prototype for chronic disease management” (Goldman and Smith, 2002). My findings, which analyze a broad range of health conditions and chronic diseases, would suggest that, insofar as the formal schooling is concerned, diabetes appears to be an exception. In the SIPP data, diabetes enters in the expected direction; that is, increases in schooling appear to reduce the incidence of severe cases of diabetes.

On the one hand, since diabetes is also associated with loss of limbs and poor vision, the diabetes result could be a plausible explanation for those findings. On the other hand, kidney problems and hypertension, which are also commonly associated with diabetes, go in the wrong direction. Further, there is no well-established connection between diabetes and speech, hearing, and back problems. An alternative explanation for the diabetes result could be that states that had higher compulsory schooling levels also promoted nutritional policies that might have reduced adult onset of diabetes. Overall, however, one conclusion that may be drawn from this table is that there is little support for the “decision-making” hypothesis.

I would also note that explanations for the link between education and health that focus on better health care access due to more financial resources (for example, from higher income and a better paying occupation) or unobserved time preferences do not appear to be consistent with these results. These explanations would likely imply that many outcomes ought to be affected, not just a few.

There are two important limitations to this analysis. First, I observe individuals only if they have survived into the 1980s and 1990s when they are anywhere between the ages of 59 and 83. This sample is almost certainly positively selected on education and health, so it is unclear to what extent they may be generalized. I suspect that because of this selection, my results are biased against finding any effects of education on improving health, making it still surprising that there are very large negative coefficients on the incidence of several negative health outcomes. Second, because

specific health conditions are only asked of those who report an activity limitation or being in fair or poor health, some individuals with a particular condition may not be captured in the analysis. Nonetheless, it may be even more meaningful to identify the effects of education on specific conditions that were severe enough to cause an activity limitation.

## Conclusion

In this article, I expand upon the growing literature that attempts to identify whether there is a causal effect of education on health. I closely examine the effects of education induced by compulsory schooling laws early in the twentieth century on long-term health, using several approaches. First, I revisit the results in Lleras-Muney (2005, 2006) by expanding the U.S. Census sample and employing a variety of robustness checks. The main finding is that the effects of education on mortality induced by changes in compulsory schooling laws are not robust to including state-specific time trends, suggesting that a causal interpretation is unwarranted.

Second, I use the SIPP to identify not only general health effects but also specific health outcomes that were induced by changes in state compulsory schooling laws to see if these outcomes correspond to our existing theories of how education affects health. The results suggest that there is a large effect of education on general health status arising from compulsory schooling laws that are robust to state time trends. However, I find that, with the important exception of diabetes, none of the other specific health conditions that are associated with education (for example, vision, hearing, speaking ability, back problems, deformities, and senility) correspond to the leading theories of how education improves health (for example, technological improvements, better decision-making, lower discount rates, higher income). This suggests that either our theories are incorrect or that the compulsory schooling laws are suspect instruments. An important caveat, however, is that the SIPP analysis uses a sample of older individuals who are almost surely positively selected on education and health. While this likely makes it more difficult to detect effects of education on improved health, it also raises questions as to how far one can generalize these results.

A few other studies have begun to implement strategies to better identify the causal effects of education on health with mixed findings. In a working paper, Clark and Royer (2007) use differences in compulsory schooling laws affecting very narrowly defined birth cohorts in the United Kingdom, combined with *individual*-level mortality data and find very small



effects of education on mortality, which are consistent with the results here. In another working paper, Deschenes (2007) uses plausibly exogenous variation based on cohort size in the U.S. and estimates a statistically significant and large effect of education on mortality using a grouped estimator. Deschenes' estimates

suggest that an additional year of schooling adds an additional year to life expectancy. Because we are still only in the early stages of our understanding of this important issue, it is important to conduct replication and extension exercises on the small number of studies that have used more credible research strategies.

---

## NOTES

<sup>1</sup>Kolata (2007).

<sup>2</sup>For example, Deaton and Paxson (2004) document that there is a strong association between education and health in the United Kingdom.

<sup>3</sup>See Lyman (2006). The National Institute on Aging is part of the National Institutes of Health.

<sup>4</sup>The results from using the Lleras-Muney (2005) instruments instead of the Goldin and Katz (2003) instruments are not very different, and are in an earlier version of this article, Mazumder (2007).

<sup>5</sup>The IPUMS are from the University of Minnesota, Minnesota Population Center.

<sup>6</sup>Lleras-Muney (2002) found no effect of compulsory schooling laws on the education levels of blacks.

<sup>7</sup>I thank Jay Bhattacharya for this suggestion. In a previous version of the article, I found very similar results using two-stage least squares for the dichotomous outcomes.

<sup>8</sup>I generally found that the IV results were larger and more significant when using the state trends than when using region of birth interacted with cohort. The ordinary least squares results were virtually identical under either specification.

<sup>9</sup>The 1990 and 1996 panels include an oversample of poorer households. The restriction to the noninstitutionalized population means

that those living in nursing homes are not included in the survey. However, more than 90 percent of the disabled and more than 80 percent of those requiring long-term care live outside of institutions; for further details, see <http://aspe.hhs.gov/daltcp/reports/rn11.htm>.

<sup>10</sup>See Johnson and Schoeni (2007) and the citations therein for a discussion of this approach.

<sup>11</sup>I pool responses from the 1984, 1990–93, and 1996 SIPPs in order to maximize sample size. Unfortunately, different criteria were used across the SIPP survey years to select the subsamples for which specific health conditions were asked. For example, in 1996 the health conditions were asked of those who reported being in fair or poor health. I found that it was important to combine all of the subsamples in all of the years in order to have enough power to identify effects. There are also an additional set of ten outcomes that are not used because they were not available in the 1984 SIPP. Experimentation with a smaller sample suggests that the conclusions are not altered by dropping these other outcomes.

<sup>12</sup>Note that these are estimates from errata that correct the previous estimates in Lleras-Muney (2005). See Mazumder (2007) for more details.

<sup>13</sup>The mean ten-year mortality rate in Lleras-Muney (2005) is 10.6 percent, so a reduction of 6.3 percentage points implies a 59 percent reduction in mortality.

<sup>14</sup>The partial  $F$  statistic rises to 9.07 when using the expanded sample.

---

## REFERENCES

- Acemoglu, D., and J. Angrist**, 2001, "How large are human capital externalities? Evidence from compulsory schooling laws," in *NBER Macroeconomics Annual 2000*, B. S. Bernanke and K. S. Rogoff (eds.), Cambridge, MA: MIT Press, pp. 9–59.
- Becker, G. S., and C. B. Mulligan**, 1997, "The endogenous determination of time preference," *Quarterly Journal of Economics*, Vol. 112, No. 3, August, pp. 729–758.
- Case, A., D. Lubotsky, and C. Paxson**, 2002, "Economic status and health in childhood: The origins of the gradient," *American Economic Review*, Vol. 92, No. 5, December, pp. 1308–1334.
- Clark, D., and H. Royer**, 2007, "The effect of education on longevity: Evidence from the United Kingdom," Case Western Reserve University, working paper.
- Cutler, D., and G. Miller**, 2005, "The role of public health improvements in health advances: The twentieth-century United States," *Demography*, Vol. 42, No. 1, February, pp. 1–22.
- Deaton, A., and C. Paxson**, 2004, "Mortality, income, and income inequality over time in Britain and the United States," in *Perspectives on the Economics of Aging*, D. A. Wise (ed.), Chicago: University of Chicago Press, pp. 247–280.

- Deschenes, O.**, 2007, "The effect of education on adult mortality: Evidence from the baby boom generation," University of California, Santa Barbara, working paper.
- Dhir, R., and J. P. Leigh**, 1997, "Schooling and frailty among seniors," *Economics of Education Review*, Vol. 16, No. 1, February, pp. 45–57.
- Glied, S., and A. Lleras-Muney**, 2003, "Health inequality, education, and medical innovation," National Bureau of Economic Review, working paper, No. 9738, June.
- Goldin, C., and L. F. Katz**, 2003, "Mass secondary schooling and the state," National Bureau of Economic Review, working paper, No. 10075, November.
- Goldman, D. P., and J. P. Smith**, 2002, "Can patient self-management help explain the SES health gradient?," *Proceedings of the National Academy of Sciences*, Vol. 99, No. 16, August 6, pp. 10929–10934.
- Grossman, M.**, 2005, "Education and nonmarket outcomes," National Bureau of Economic Review, working paper, No. 11582, August.
- Gunderson, G. W.**, 1971, "The National School Lunch Program: Background and development," U.S. Department of Agriculture, Food and Nutrition Service, report, available at [www.fns.usda.gov/cnd/Lunch/AboutLunch/ProgramHistory.htm](http://www.fns.usda.gov/cnd/Lunch/AboutLunch/ProgramHistory.htm).
- Hunter, R.**, 1904, *Poverty*, New York: Macmillan.
- Johnson, R. C., and R. F. Schoeni**, 2007, "The influence of early-life events on human capital, health status, and labor market outcomes over the life course," University of California, Berkeley, Institute for Research on Labor and Employment, working paper, No. iirwps-140-07, January 2.
- Kitagawa, E. M., and P. M. Hauser**, 1973, *Differential Mortality in the United States: A Study in Socio-economic Epidemiology*, Cambridge, MA: Harvard University Press.
- Kolata, G.**, 2007, "A surprising secret to a long life: Stay in school," *New York Times*, January 3, available at [www.nytimes.com/2007/01/03/health/03aging.html](http://www.nytimes.com/2007/01/03/health/03aging.html).
- Lleras-Muney, A.**, 2006, "Erratum: The relationship between education and adult mortality in the United States," *Review of Economic Studies*, Vol. 73, No. 3, p. 847.
- \_\_\_\_\_, 2005, "The relationship between education and adult mortality in the United States," *Review of Economic Studies*, Vol. 72, No. 1, pp. 189–221.
- \_\_\_\_\_, 2002, "Were compulsory attendance and child labor laws effective? An analysis from 1915 to 1939," *Journal of Law and Economics*, Vol. 45, No. 2, part 1, October, pp. 401–435.
- Lyman, R.**, 2006, "Census report foresees no crisis over aging generation's health," *New York Times*, March 10, available at [www.nytimes.com/2006/03/10/national/10aging.html](http://www.nytimes.com/2006/03/10/national/10aging.html).
- Mazumder, B.**, 2007, "How did schooling laws improve long-term health and lower mortality?," Federal Reserve Bank of Chicago, working paper, No. WP-2006-23, revised January 24, 2007.
- Marmot, M. G.**, 1994, "Social differences in health within and between populations," *Daedalus*, Vol. 123, No. 4, pp. 197–216.
- National Institutes of Health**, 2003, "Pathways linking education to health," report, Bethesda, MD, No. RFA OB-03-001, January 8, available at <http://grants1.nih.gov/grants/guide/rfa-files/RFA-OB-03-001.html>.
- Rivers, D., and Q. H. Vuong**, 1988, "Limited information estimators and exogeneity tests for simultaneous probit models," *Journal of Econometrics*, Vol. 39, No. 3, November, pp. 347–366.

11/12/10

**ERRATUM, Corrected Table 3: New Estimates of effects of education on mortality**

		IV	Number of observations
<b>A. 1960 - 1980</b>			
1960-1980 1%			
No age controls, region X cohort	-0.036 (0.004)	-0.072 (0.025)	4792
1960 1%, 1970 2%, and 1980 5%:			
No age controls, region × cohort	-0.045 (0.004)	-0.045 (0.024)	4797
With age cubic, region × cohort	-0.039 (0.004)	-0.047 (0.024)	4797
With age cubic × Census year, region × cohort	-0.039 (0.004)	-0.047 (0.024)	4797
With age cubic × Census year, state × cohort trend	-0.040 (0.004)	0.003 (0.038)	4797
<b>B. 1960 - 2000</b>			
1960 1%, 1970 2%, and 1980–2000 5%:			
With age cubic × Census year	-0.034 (0.003)	-0.029 (0.015)	8636
With age cubic × Census year, state × cohort trend	-0.035 (0.003)	0.006 (0.031)	8636
<b>C. 1960–2000, by Census year</b>			
1960 1%, 1970 2%, and 1980–2000 5% with age cubic:			
Estimated effect for 1960–70	-0.025 (0.006)	-0.081 (0.052)	2397
Estimated effect for 1970–80	-0.061 (0.005)	-0.023 (0.033)	2400
Estimated effect for 1980–90	-0.043 (0.004)	0.023 (0.029)	2399
Estimated effect for 1990–2000	-0.012 (0.005)	0.027 (0.039)	1440
<b>D. 1960–2000, by age</b>			
1960 1%, 1970 2%, and 1980–2000 5% with age cubic × Census year:			
35–54 year olds	-0.017 (0.005)	-0.064 (0.036)	2879
55–64 year olds	-0.039 (0.005)	0.063 (0.053)	2398
65–89 year olds	-0.031 (0.003)	-0.052 (0.022)	3359
<b>E. 1960–2000, by cohort</b>			
1960 1%, 1970 2%, and 1980–2000 5% with age cubic × Census year:			
Cohorts born in 1901–12	-0.019 (0.004)	-0.200 (0.124)	3644
Cohorts born in 1913–25	-0.017 (0.004)	0.029 (0.023)	4992

Notes: WLS means weighted least squares. IV means instrumental variables. The dependent variable is the ten-year mortality rate; table entries are the coefficient on education. All specifications include year dummies, cohort dummies, state of birth dummies, region of birth interacted with cohort, and an intercept (except for panel A, fifth row, and panel B, second row). Estimates are weighted using the number of observations in the cell in the base year. Standard errors, shown in parentheses, are clustered at the state of birth and cohort level.

# How do EITC recipients spend their refunds?

Andrew Goodman-Bacon and Leslie McGranahan

## Introduction and summary

The earned income tax credit (EITC) is one of the largest sources of public support for lower-income working families in the U.S. The EITC operates as a tax credit that serves to offset the payroll taxes and supplement the wages of low-income workers. For tax year 2004, the EITC transferred over \$40 billion to 22 million recipient families (U.S. Internal Revenue Service, 2006b). Nearly 90 percent of program expenditures come in the form of tax refunds; the remaining 10 percent serve to reduce tax liability. While other income support programs distribute benefits fairly evenly across the calendar year, EITC payments are concentrated in February and March when tax refunds are received. Because the EITC makes one relatively large payment per year, it may provide low-income, credit-constrained households with a rare opportunity to make important big-ticket purchases.

Research on the EITC has tended to focus on the important labor supply effects of the program (Eissa and Liebman, 1996; Meyer and Rosenbaum, 2001; and Grogger, 2003). Relatively little is known about how recipient households actually use EITC refunds. In this article, we use data from the U.S. Bureau of Labor Statistics' *Consumer Expenditure Survey* (CES) over the period 1997–2006 to investigate how households spend EITC refunds.<sup>1</sup> Following the methodology of Barrow and McGranahan (2000), we rely on the particular timing of EITC payouts to identify the effects of the credit on expenditures. Barrow and McGranahan found that the EITC has a larger effect on spending on durable goods than on nondurable goods. In this article, we are particularly interested in determining what items within the durables and nondurables categories are purchased using the credit and whether these expenditures reinforce the EITC's prowork and prochild goals. Our primary finding is that recipient household spending in response to EITC payments is

concentrated in vehicle purchases and transportation spending. Given the crucial link between transportation and access to jobs, we believe this finding is consistent with the EITC's goals. In the next section, we present a brief history of the EITC and the key features of the program. We then review prior research on the uses of the EITC by recipient families. Next, we introduce the CES data and the methodology we use to investigate the data. Finally, we present our results and discuss their implications.

## History and structure of the EITC

Congress created the EITC in 1975 to offset payroll taxes paid by low-income workers with children. The credit is structured as a supplement to earned income equaling a percentage of earnings up to a specific threshold (the "phase-in" range), at which point the credit amount stays constant for an additional amount of earnings (the "plateau" range). Then this maximum credit is reduced by a given percentage of earnings until it equals zero (the "phase-out" range). Income thresholds, the phase-in and phase-out rates, and, therefore, the credit amount also vary by the number of qualified children in a household and by marital status; and all these factors have varied over time.<sup>2</sup> Figure 1 graphs the EITC program parameters for selected years. The program is implemented as a part of the tax code, and recipients must file taxes in order to apply for the program. For tax year 2006, a single mother with two children earning between \$11,340 and \$14,810 would have received the maximum credit of \$4,536.

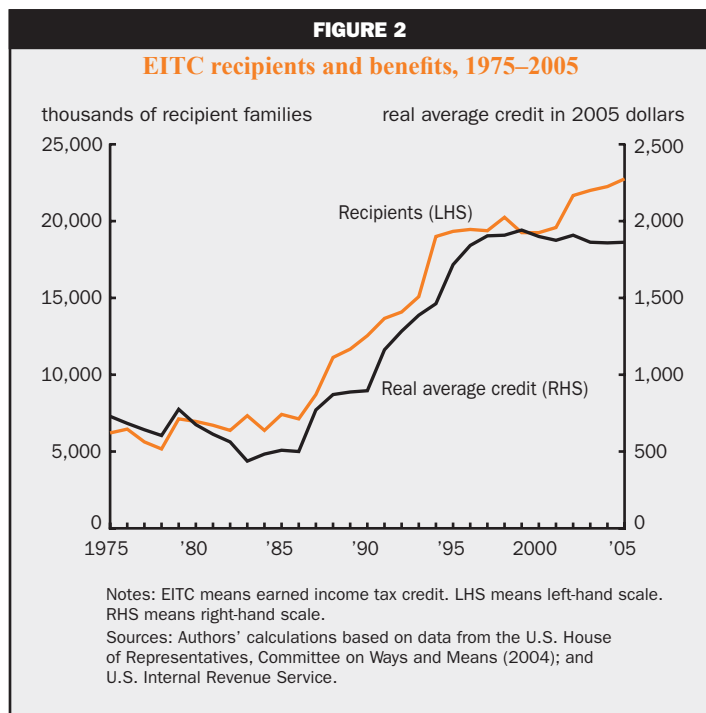
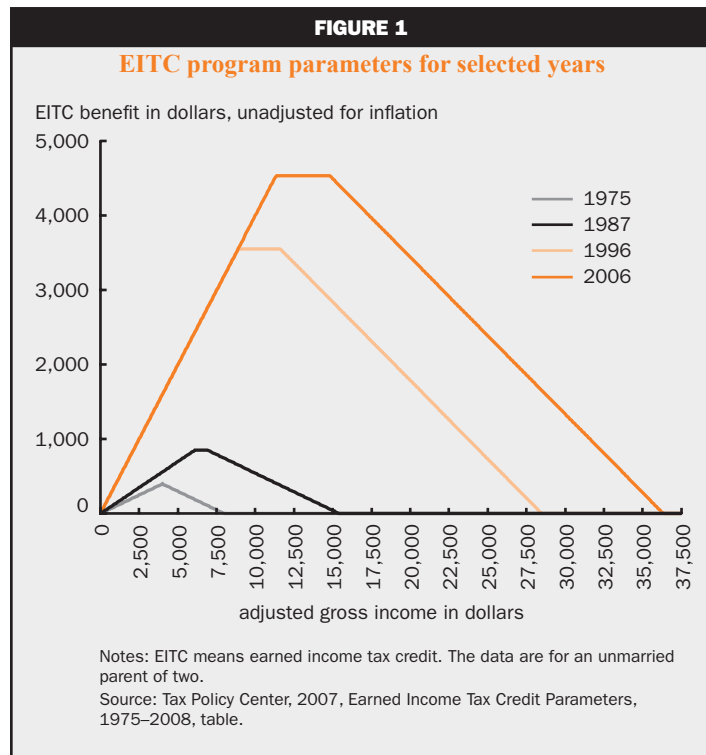
*Andrew Goodman-Bacon is currently a graduate student in economics at the University of Michigan and a former associate economist at the Federal Reserve Bank of Chicago. Leslie McGranahan is an economist in the Economic Research Department at the Federal Reserve Bank of Chicago. The authors thank Lisa Barrow, Eric French, and Anna Paulson for helpful comments.*



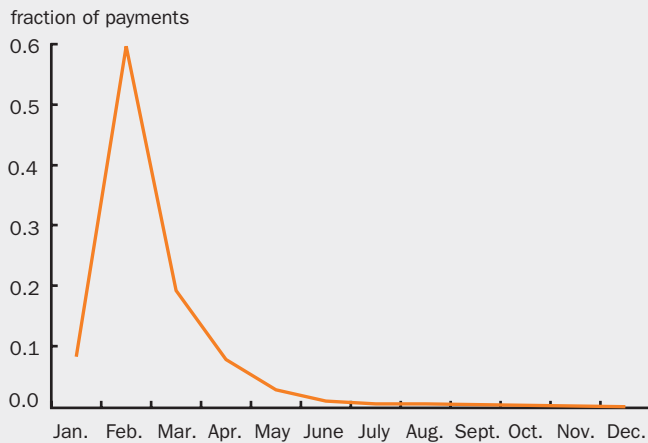
The EITC began as a small program, but its generosity and coverage have expanded frequently in its 30-year history as is shown in figures 1 and 2. Particularly large expansions enacted in 1986 and 1993 led to rapid program growth. In 1994, childless families started to receive a small credit. In 1975, the EITC represented 3.1 percent of federal means-tested transfers and 9.7 percent of federal means-tested cash transfers; by 2002, these proportions had increased by three times and four and a half times, respectively, and the EITC was the second largest means-tested cash transfer program behind Supplemental Security Income (SSI). In figure 2, we graph the average credit and number of recipient families by year. As the figure shows, the size of the EITC was relatively constant in its first decade, but between 1986 and 2005, both the number of recipient families and the real average credit amount grew by more than three times, increasing real federal expenditures on the program by almost 12 times. In 1986, just over 7 million families received earned income tax credits averaging \$501 in 2005 dollars. By 2002, over 20 million families received credits averaging \$1,911 in 2005 dollars (U.S. House of Representatives, Committee on Ways and Means, 2004).

Unlike other transfer programs that have monthly benefits, the EITC pays out a lump sum once per year. The EITC does permit recipients to receive some portion of payments monthly prior to tax filing in the form of the advance earned income tax credit, but in 2004 only 0.6 percent of recipient households received any credit in this manner, representing just 0.2 percent of payments (U.S. Internal Revenue Service, 2006b).

Figure 3 shows the distribution of refundable EITC payments from the U.S. Internal Revenue Service by month for 2005—a year with a payment pattern typical of recent years. As the figure shows, nearly all EITC payments are made in February and March, and most of these come in February. The modal month of EITC payments has changed over time, but since 1997 more payments have been made in February than in any other month.



This pattern is a result of the timing of tax filing. Taxes can be filed anytime after W-2s (employee wage report forms) are received (by January 31), and refunds are received within six weeks.<sup>3</sup>

**FIGURE 3****Fraction of EITC payments, by month, 2005**

Note: EITC means earned income tax credit.

Source: Authors' calculations based on data from the U.S. Department of the Treasury, 2005–06, *Monthly Treasury Statement of Receipts and Outlays of the United States Government*, various issues.

The lump sum payment structure also means that EITC refunds represent a relatively large share of recipients' income in the month when they are received. For tax year 2004, the average EITC refund for recipient families with children was \$2,113, or 12 percent of their annual average adjusted gross income (AGI) of \$16,981. Assuming income was earned evenly across the calendar year, the average recipient household's income would be approximately two and a half times its usual monthly value in the month when the EITC payment was received.<sup>4</sup>

For comparison, the mean overpayment refund for non-EITC recipients in tax year 2004 was \$1,692, or 2.9 percent of annual average AGI among nonrecipients.<sup>5</sup> Overpayment refunds are less concentrated in the first quarter of the year than EITC refunds. While 87 percent of EITC refunded dollars for 2004 were distributed in the first quarter, 47 percent of non-EITC refunded dollars were distributed in the first quarter, and an additional 42 percent were distributed in the second quarter (U.S. Internal Revenue Service, 2006c). It is worth noting that the *Consumer Expenditure Survey*, the data set used for our analysis, provides additional evidence to show that EITC refunds are concentrated earlier in the year than other tax refunds. Among families who made an expenditure on "accounting services," including tax preparation, 43 percent of EITC eligible families did so in January or February, versus 29 percent of noneligible families.

The one-time EITC average refund of \$2,113 among families with children in 2004 is also large when compared with the average monthly payments to recipient families in other transfer programs in 2004, such as SSI (\$429); Temporary Assistance to Needy Families, or TANF (\$397); the Food Stamp Program (\$200); and unemployment insurance (\$1,141).<sup>6</sup>

### Use of EITC refunds

The majority of research on the EITC and expenditure patterns has relied on surveys of EITC recipients about how they spent or planned to spend refunds. The consensus from these surveys is that the primary use of EITC refunds is to pay bills. Sixty-three percent of respondents in a survey of participants in the University of Georgia's Consumer Financial Literacy Program reported that they planned to use most of their refund to pay or catch up on bills or debts (Linnenbrink et al., 2006).

Similarly, 44 percent of mothers in a study tracking the well-being of rural families indicated that they used their refund to pay bills (Mammen and Lawrence, 2006). Using surveys of free tax preparation clients in Chicago, Smeeding, Phillips, and O'Connor (2000) report that tax filers who anticipate an EITC refund most often plan to use it to pay bills. These studies also find that recipients used their refunds to purchase or repair cars and buy other durables, such as home furnishings. Some families also report buying children's clothing and going on vacation. Very few families planned to save their refund for a rainy day or for retirement.

In contrast to these studies, Barrow and McGranahan (2000) use the nationally representative *Consumer Expenditure Survey* to investigate expenditure uses of EITC refunds. They rely on the unique seasonal pattern of EITC refunds to determine whether EITC eligible households have expenditure patterns that differ from those of noneligible households. They find that EITC eligible households have higher expenditures on durable goods in February, the modal month of EITC receipt, relative to noneligible households. They attribute this increased spending on durables to the EITC. Barrow and McGranahan do not measure health care, housing, or utility expenditures, so they do not measure much of what other studies categorize as "bills."

Here we use CES data over the period 1997–2006 to build upon the work of Barrow and McGranahan (2000). We investigate on which goods, particularly

within the durable goods category, the EITC recipient households spend more. We also look at both the extensive and intensive margin of expenditure. In other words, we ask both whether households are more likely to make any expenditure and whether they make larger expenditures, given that they make a purchase.

We focus on those goods that have been identified in the literature as either those that recipients report that they plan to purchase or those that further the EITC program's goals of "strengthen[ing] the incentive to work," "help[ing] low-wage working families make ends meet," and promoting the well-being of children (Frost, 1993). Vehicle expenditures fall into both of these categories. They have been mentioned by recipients as an intended use of the EITC credit and are particularly supportive to work. According to a Brookings Institute report, 88 percent of low-income Americans commute in a personal vehicle (Blumenberg and Waller, 2003). In fact, other antipoverty and income support programs explicitly recognize the link between car ownership and employment through more lenient limits on cars than on other forms of assets. For example, the federal SSI program exempts one vehicle from its resource limit. Similarly, most states exclude the value of one or more vehicles from resource limits used to determine eligibility for the Food Stamp Program and TANF Program. In addition to vehicles, we focus on expenditures on household furnishings and home electronics, as well as on children's clothing. We do not look at bill paying because the nature of the CES data precludes such an analysis.

Our primary contribution is to provide evidence on detailed actual expenditures, using nationally representative survey data. Time-series variation in EITC payments over the year and cross-sectional variation in imputed eligibility allow us to identify the EITC's impact. Similar to Barrow and McGranahan (2000), we find that receiving EITC refunds increases household expenditures on both durable and nondurable goods, but more so for durables. Eligible households are more likely both to purchase big-ticket items in February and to spend more on them, given that they make any expenditure. Within durables, the strongest patterns are found for vehicles, confirming the responses given in surveys. Eligible households also spend slightly more on all other major subcategories of durables—household goods, appliances, and home electronics. Within nondurables, the strongest patterns are found for transportation expenses, such as car repairs.<sup>7</sup>

## Data

We create a monthly household-level data set of expenditure, income, and family structure, using the

CES's interview survey data covering the period 1997–2006. Households, which are called consumer units (CUs) in the data, are interviewed five times for the survey.<sup>8</sup> The first interview provides baseline asset information. The second through fifth interviews cover detailed expenditure information for the three months prior to the interview date. These interviews occur three months apart. As a result, in the absence of attrition, a full year of expenditure data is collected for each household. Households enter and exit the survey each month. Information on income in the 12 months leading up to the survey date is collected in the second and fifth interviews. Demographic information is updated every interview.

We begin with the 1997 data because February has been the modal month of EITC payouts since 1997. This consistency in payments across time allows us to focus on the February expenditures of recipient households. In most years prior to 1997, March was the modal month of EITC payments.<sup>9</sup>

We consider a CU to belong to the calendar year in which we observe February expenditure (or would have observed it if the household had responded). Since 1997, this is when the CU is most likely to have received the previous tax year's EITC refund payment. Therefore, data over the period 1997–2006 allow us to consider EITC policies in place during tax years 1996–2005. The average number of observations in our 120 month-year cells is 4,888, and in total we have 589,568 observations.

Information on EITC receipt is not provided in the CES, so we use the income and family structure variables to impute EITC eligibility and the magnitude of EITC payments. Because of our reliance on the income data, we delete those with incomplete income reports from the analysis. We assume all households without children are not eligible for the EITC despite the small credit for childless families that has been available since 1994.<sup>10</sup> The CUs may contain more than one tax filing unit (TU). We impute EITC payments and eligibility for each TU within the CU and combine these to determine CU eligibility and EITC amount. Ideally, we would observe the income and family structure of each TU for the year preceding their February interview. However, we lack information on TU composition and on tax year income. To generate our best guess of income for the year preceding the February interview, we use the income information in the second and fifth interviews. For some individuals, our best guess of tax year income is the reported income from the second interview; for others, we compute a weighted average of the two income reports where

the weights depend on the number of months for which the year covered by the income report and tax year overlap.

To assign adults to TUs and generate TU income, we use sex, marital status, relationship to reference person, and individual income information. To assign children to TUs for the purpose of the EITC computation, we use the EITC eligibility rules in place during the year before their February interview. Before 2001, EITC rules assigned all qualifying children in a *family* to the TU with the highest income, but since 2001, families have been free to choose which TU claimed qualifying children. Thus, before 2001 we give all children to the highest-income TU, and after 2001 we give all qualifying children to the TU for which they generate the largest EITC refund.<sup>11</sup>

Because of this imputation, we are measuring EITC eligibility rather than EITC receipt. Two issues may affect the accuracy of these imputations. First, some households that are eligible for the EITC may not take it up. According to a study by the U.S. Government Accounting Office (2001), approximately 85 percent of eligible households with children participate in the EITC program. Second, we may be incorrectly imputing that eligible households are ineligible or that ineligible households are eligible because either child or income information is incorrect in the CES. There is some underreporting of income in the CES, so we may be assigning eligibility to some households that

are in fact beyond the maximum income for EITC receipt. We also may be assigning some children to an incorrect TU. These issues make it harder for us to find an effect of the EITC on consumption. As a result, our estimates represent a lower bound on the effect of the EITC on recipient consumption patterns.

Table 1 gives variable means for the demographic, income, and EITC variables for all families and by imputed EITC eligibility. In the full sample, 13 percent of household-months (shown as 0.13 in the first column, fourth row of table 1) were eligible for an average credit of \$2,116 in the February in which we observed them. These percentages and values change over time in keeping with the changes in eligibility and refund amounts presented in figure 2 (p. 18). When we compare the EITC eligible and noneligible populations, we find differences that are consistent with the program rules. For example, EITC eligible households earn approximately 60 percent of what noneligible households earn on average, and have more children. In addition, EITC eligible households are also less likely to have a white household head, are more likely to be headed by a single parent, and are less educated than noneligible households. These additional findings are not related explicitly to the program rules, but result from patterns of earnings in the U.S., and are consistent with the attributes of participants in other income support programs.

**TABLE 1**

**Summary statistics**

	All	Non-EITC	EITC
Median real income (2004 dollars)	32,346	36,590	22,548
Mean real income (2004 dollars)	44,130	46,468	28,599
EITC amount (2004 dollars)	277	—	2,116
EITC eligible	0.13	0.00	1.00
Number of children	0.71	0.52	1.97
White household head	0.84	0.85	0.75
Household head's highest educational attainment:			
Some high school	0.13	0.12	0.22
High school diploma	0.25	0.24	0.34
Some college	0.20	0.20	0.23
College degree	0.42	0.45	0.22
Family type:			
Husband, wife, and own kids	0.27	0.25	0.37
Single parent	0.06	0.03	0.25
Single person	0.28	0.32	0.00
Other family type	0.39	0.40	0.39
Observations (family months)	589,568	512,405	77,163
Observations (distinct families)	59,595	51,824	7,771

Note: EITC means earned income tax credit.  
Source: Authors' calculations based on data from the U.S. Bureau of Labor Statistics, *Consumer Expenditure Survey*.

Our next goal is to generate monthly expenditure data. We combine all available interviews for each CU. Sixty-three percent of CUs have 12 months of data, and the average CU has 9.9 months of data. The CES contains very detailed information on expenditures, which we distill into durable goods and nondurable goods, as well as subcategories of those groups. Durable goods includes household goods (such as furniture, linens, and carpets); appliances (such as dishwashers, silverware, and kitchen electronics); electronics (such as televisions and computers); and new and used vehicle purchases. Nondurables includes food, alcohol, and tobacco; apparel; trips (out-of-town travel and expenditure while traveling); transportation expenses (except vehicle purchases); entertainment; child support, alimony, and charity; and pensions, insurance, and social security payments. We do not measure expenditure on items that we do not consider to be durable or nondurable goods. In particular, we exclude utilities, rent, education expenses, and health care. These obligations may be difficult for households to alter on a month-to-month basis. In addition, the rent and utility variables reported on the survey capture the amount owed in a given month rather than the amount paid, making it impossible to assess whether households are spending money to catch up on overdue payments or prepay obligations.<sup>12</sup>

Table 2 provides summary statistics on expenditures in all of our categories as calculated from the CES. It provides three different measures of expenditure for each category. The first set of three columns presents expenditure that occurs on the goods category in the average month as a percent of total annual expenditures on durable and nondurable goods. The entry for durable goods in the first column indicates that in the average month, a household spends 1.5 percent of its total annual durable and nondurable goods expenditures on durable goods. The second set of three columns reports the probability that a household makes any expenditure in a category in an average month. In the average month, 84.5 percent of households purchase a durable good. The third set of three columns reports the proportion of total annual expenditure for durable and nondurable goods in that category in a month, given that some expenditure was made. Among households purchasing durables in a given month, the average household spends 1.8 percent of total annual durable and nondurable goods expenditures on durables. Table 3 reports the average dollar amount (in 2004 dollars based on the Personal Consumption Expenditures deflator) spent per month conditional on expenditure.

As seen in table 2, average monthly expenditure shares are fairly consistent for EITC and non-EITC families with a few exceptions. The EITC families spend a high share on food and on children's clothing. The higher expenditure share on food is consistent with the general finding that food expenditure shares are higher for lower-income households in the U.S. The higher expenditure share on children's clothing arises from our restriction that all EITC eligible households have children, while many noneligible households do not. From the second group of columns in table 2, we observe that EITC families are generally less likely than non-EITC families to make expenditures in almost every category in an average month. As shown in table 3, in dollar terms, conditional on nonzero expenditure, EITC families spend less on everything except for tobacco, food, and gasoline. Our analysis continues by examining the effect of EITC eligibility on spending in the nondurables category and the nondurable goods subcategories of children's clothing and transportation, and then we focus our analysis on durable goods expenditures and specifically on expenditures for vehicles and consumer electronics.

## Methodology

We measure expenditure by household  $i$  in month  $t$  on category  $j$  in three ways: the proportion of annual

expenditure in each month  $\left( \frac{X_{it}^j}{X_{i,Annual}^j} \right)$ , the probability

of making any expenditure  $\left( P(X_{it}^j > 0) \right)$ , and the proportion of annual expenditure conditional on making

an expenditure  $\left( \frac{X_{it}^j}{X_{i,Annual}^j} \mid X_{it}^j > 0 \right)$ .<sup>13</sup>

We estimate clustered probit models for the discrete measure of expenditure and generalized least squares (GLS) regression models for the expenditure proportion variables. Letting  $X$  be one of the three dependent variables, we estimate the following equation:

$$1) \quad X_{it}^j = \alpha + \gamma_t M_t + \phi EITC_i + \lambda_t (EITC_i \times M_t) + \beta C_i + \varepsilon_{it},$$

where  $M$  is a vector of month dummies,  $EITC$  is a dummy variable equal to 1 if the household is imputed to be EITC eligible, and  $C$  is a vector of household-level controls—year of first quarter interview; income, race, sex, and education of household head; family size; number of children; family type; and region (all rural households are the omitted “region”).



TABLE 2

## Expenditure patterns, by expenditure category and EITC eligibility

	Monthly expenditure/ annual expenditure			Probability of expenditure			Monthly expenditure/ annual expenditure, conditional on nonzero expenditure		
	All	Non-EITC	EITC	All	Non-EITC	EITC	All	Non-EITC	EITC
<b>Total</b>	<b>0.084</b>	<b>0.084</b>	<b>0.083</b>	<b>1.000</b>	<b>1.000</b>	<b>1.000</b>	<b>0.084</b>	<b>0.084</b>	<b>0.083</b>
<b>Durable goods</b>	<b>0.015</b>	<b>0.015</b>	<b>0.015</b>	<b>0.845</b>	<b>0.848</b>	<b>0.822</b>	<b>0.018</b>	<b>0.018</b>	<b>0.019</b>
Household goods	0.003	0.003	0.002	0.285	0.291	0.248	0.010	0.010	0.009
Furniture	0.001	0.001	0.001	0.048	0.048	0.048	0.027	0.028	0.024
Drapes, linens, and floor coverings	0.000	0.001	0.000	0.098	0.098	0.094	0.005	0.005	0.004
Miscellaneous household equipment	0.001	0.001	0.001	0.210	0.216	0.169	0.005	0.005	0.004
Appliances	0.001	0.001	0.001	0.101	0.102	0.100	0.009	0.009	0.008
Major appliances	0.001	0.001	0.001	0.031	0.031	0.032	0.022	0.023	0.019
Minor appliances	0.000	0.000	0.000	0.076	0.076	0.074	0.003	0.003	0.003
Electronics	0.004	0.004	0.004	0.809	0.813	0.783	0.005	0.005	0.005
Vehicle purchases	0.007	0.007	0.008	0.024	0.022	0.034	0.302	0.316	0.244
<b>Nondurables</b>	<b>0.068</b>	<b>0.068</b>	<b>0.068</b>	<b>0.999</b>	<b>0.999</b>	<b>1.000</b>	<b>0.068</b>	<b>0.068</b>	<b>0.068</b>
Food, alcohol, and tobacco	0.030	0.030	0.034	0.997	0.997	0.998	0.030	0.030	0.035
Food	0.023	0.022	0.028	0.992	0.991	0.994	0.023	0.022	0.028
Alcohol	0.001	0.001	0.001	0.350	0.363	0.265	0.002	0.002	0.002
Tobacco	0.002	0.002	0.002	0.257	0.244	0.341	0.007	0.007	0.007
Food away from home	0.005	0.005	0.004	0.808	0.813	0.772	0.006	0.007	0.005
Apparel	0.006	0.005	0.007	0.647	0.643	0.674	0.009	0.009	0.010
Trips	0.003	0.003	0.002	0.186	0.196	0.119	0.017	0.017	0.014
Transportation	0.017	0.016	0.017	0.939	0.938	0.949	0.018	0.018	0.018
Gasoline	0.006	0.006	0.007	0.893	0.894	0.891	0.007	0.007	0.008
Other vehicle expenses	0.010	0.010	0.009	0.671	0.673	0.660	0.014	0.015	0.014
Public transportation	0.001	0.001	0.001	0.134	0.133	0.135	0.005	0.005	0.005
Entertainment	0.006	0.006	0.005	0.901	0.908	0.858	0.007	0.007	0.005
Fees, admissions, toys, and sports	0.004	0.004	0.003	0.670	0.675	0.635	0.006	0.006	0.005
Personal care services	0.001	0.001	0.001	0.734	0.749	0.635	0.002	0.002	0.002
Reading	0.001	0.001	0.000	0.587	0.611	0.428	0.001	0.001	0.001
Other nondurables	0.006	0.007	0.004	0.534	0.554	0.404	0.012	0.012	0.009
Child support, alimony, and charity	0.005	0.005	0.002	0.456	0.475	0.330	0.010	0.011	0.007
Pensions, insurance, and social security	0.002	0.002	0.001	0.189	0.196	0.142	0.009	0.009	0.008
Children's clothing	0.001	0.001	0.003	0.199	0.171	0.386	0.006	0.005	0.008
Children's clothing only among families with children	0.003	0.002	0.003	0.411	0.425	0.386	0.006	0.006	0.008

Notes: EITC means earned income tax credit. For each column, the subcategories may not total because of rounding. Children's clothing is a portion of the apparel subcategory.  
Source: Authors' calculations based on data from the U.S. Bureau of Labor Statistics, *Consumer Expenditure Survey*.

TABLE 3

## Expenditure amounts, by EITC eligibility, conditional on expenditure

	All	Non-EITC	EITC
	(-----2004 dollars-----)		
<b>Total</b>	<b>1,788.78</b>	<b>1,822.02</b>	<b>1,568.02</b>
<b>Durable goods</b>	<b>475.38</b>	<b>484.59</b>	<b>414.21</b>
Household goods	73.36	77.29	47.30
Furniture	34.27	35.81	24.02
Drapes, linens, and floor coverings	12.24	12.94	7.60
Miscellaneous household equipment	26.85	28.54	15.68
Appliances	20.24	20.97	15.38
Major appliances	14.97	15.52	11.32
Minor appliances	5.27	5.45	4.07
Electronics	79.38	80.87	69.47
Vehicle purchases	302.40	305.47	282.05
<b>Nondurables</b>	<b>1,313.40</b>	<b>1,337.43</b>	<b>1,153.81</b>
Food, alcohol, and tobacco	491.73	487.31	521.08
Food	349.55	341.46	403.29
Alcohol	15.00	15.66	10.60
Tobacco	26.19	24.80	35.42
Food away from home	101.00	105.40	71.77
Apparel	119.46	119.98	115.99
Trips	77.85	83.91	37.57
Transportation	323.12	326.16	302.99
Gasoline	109.67	108.25	119.06
Other vehicle expenses	201.27	205.32	174.40
Public transportation	12.19	12.59	9.53
Entertainment	144.07	151.83	92.53
Fees, admissions, toys, and sports	104.97	110.84	65.97
Personal care services	25.22	26.08	19.52
Reading	13.88	14.91	7.04
Other nondurables	157.15	168.23	83.63
Child support, alimony, and charity	119.03	127.82	60.63
Pensions, insurance, and social security	38.13	40.40	23.00
Children's clothing	25.20	22.12	45.68
Children's clothing only among families with children	55.31	60.62	45.68

Notes: EITC means earned income tax credit. For each column, the subcategories may not total because of rounding. Children's clothing is a portion of the apparel subcategory.

Source: Authors' calculations based on data from the U.S. Bureau of Labor Statistics, *Consumer Expenditure Survey*.

We allow for correlation among errors ( $\varepsilon$ ) within a consumer unit over time.

The coefficients in the vector  $\gamma_t$  measure the common seasonal pattern of expenditure for all households relative to September (the omitted month). For the equation measuring the percentage of total expenditure,  $\gamma_t$  indicates the fraction of total expenditure on good  $j$  in month  $t$  relative to the fraction of total expenditure in September. The coefficient  $\phi$  measures the constant difference in the fraction of expenditures between EITC eligible and noneligible households. Our coefficients of interest are the elements of the vector  $\lambda_t$ , which measure the monthly differences in expenditure (the different seasonality) between eligible and noneligible households. If all households perfectly smoothed their consumption across months,  $\gamma_t$  would be 0 and

the difference in expenditures between EITC eligible and noneligible households would be constant and entirely captured by  $\phi$ . We interpret the coefficient on the  $EITC \times February$  interaction ( $\lambda_{Feb}$ ) as an indicator of whether the EITC changes the expenditure patterns of recipients and report  $p$  values for a test of the hypothesis that  $\lambda_{Feb} = 0$ .

Our identification strategy relies on two sources of variation: cross-sectional differences in eligibility and the particular timing of EITC refunds. We have no reason to believe, a priori, that unobserved factors such as prices or preferences influence February expenditure among low-income, working families with children differently than other families.<sup>14</sup> Thus, we feel confident interpreting our  $\lambda_{Feb}$  as the impact of the EITC.

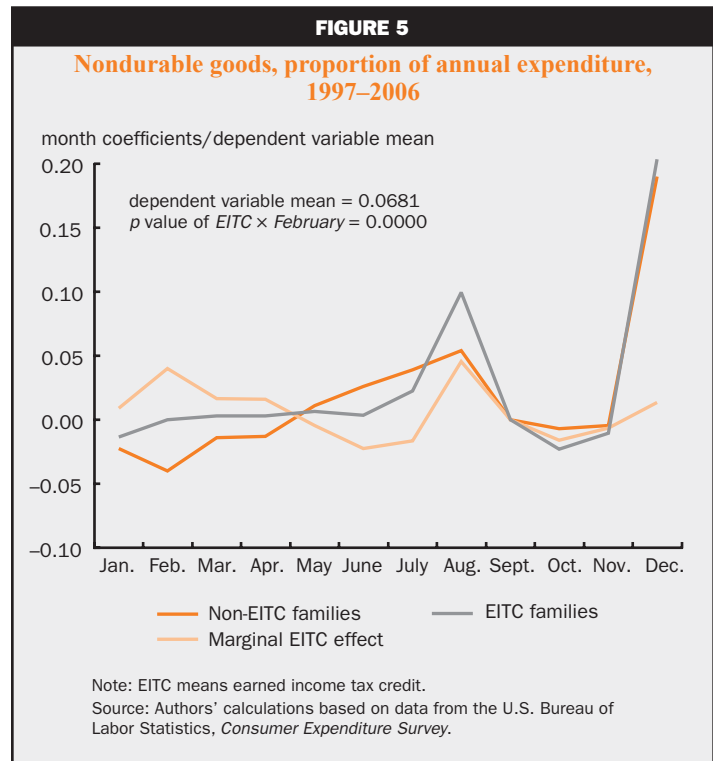
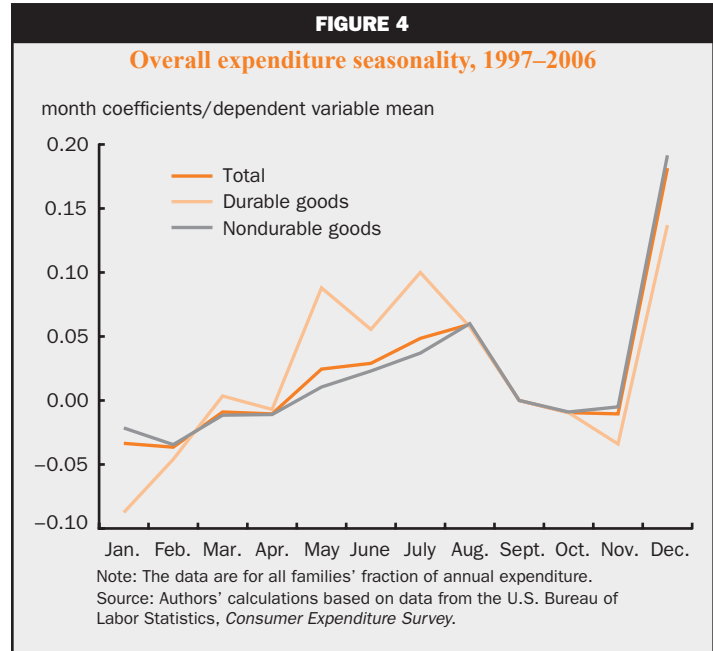
## Results

Figure 4 shows overall expenditure seasonality relative to September. There are a number of notable patterns in the data. High expenditure in December due to the holiday season dominates expenditure patterns. We also observe high durable goods expenditures in the summer months when many individuals buy cars and household items. There is also an increase in nondurable goods expenditures in August in part because of back-to-school shopping. Finally, expenditure is low in February, the shortest month of the year.

Table 4 presents estimates of  $\lambda_{Feb}$  and the associated  $p$  value for the two continuous specifications of equation 1 and marginal effects based on  $\lambda_{Feb}$  and the associated  $p$  value for the probit model. We present these results for total durable and nondurable expenditure and for numerous subcategories of expenditure. Figures 5–10 graph the coefficients  $\gamma_t$ ,  $\lambda_p$ , and  $(\gamma_t + \lambda_p)$ —labeled “Non-EITC families,” “Marginal EITC effect,” and “EITC families,” respectively, in the legend—for the three different specifications of equation 1 and for selected expenditure categories. Since we omit September and do not graph  $\phi$ , the “Non-EITC families” and “EITC families” lines represent deviations from their respective September expenditure measures. “Marginal EITC effect” is the difference between these two lines. In order to facilitate comparison between goods, for the continuous variables, the figures scale the estimated coefficients by the dependent variable mean (the average monthly expenditure on that good). For the probit model, we divide the coefficient by the estimated probability of expenditure. The denominators are listed in each figure panel, along with the  $p$  value for a test of the hypothesis that  $\lambda_{Feb} = 0$ . If  $\lambda_{Feb} = 0$ , then we cannot reject the hypothesis that the EITC does not affect expenditure on that good.

### Nondurable goods

Figure 5 depicts seasonal expenditure patterns for nondurable goods expenditures by EITC eligibility status.



As shown in the figure, we find a small, but statistically significant and positive, February effect on unconditional expenditures for EITC families ( $p$  value = 0.000). While noneligible families spend about 4 percent less on nondurables in February

**TABLE 4**  
**Effects of EITC eligibility on February expenditures**

	Unconditional expenditure		Conditional expenditure		Discrete expenditure	
	Feb. coefficient	$\rho$ value	Feb. coefficient	$\rho$ value	Feb. marginal effect	$\rho$ value
<b>Total</b>	<b>0.0067</b>	<b>0.0000</b>	<b>0.0067</b>	<b>0.0000</b>		
<b>Durable goods</b>	<b>0.0039</b>	<b>0.0004</b>	<b>0.0043</b>	<b>0.0012</b>	<b>0.0144</b>	<b>0.0023</b>
Household goods	0.0009	0.0001	0.0024	0.0087	0.0312	0.0002
Furniture	0.0008	0.0000	0.0050	0.1133	0.0195	0.0000
Drapes, linens, and floor coverings	0.0000	0.7757	-0.0008	0.2527	0.0205	0.0004
Miscellaneous household equipment	0.0001	0.2578	0.0001	0.9176	0.0238	0.0020
Appliances	0.0003	0.0125	0.0003	0.8035	0.0304	0.0000
Major appliances	0.0001	0.1710	-0.0021	0.4006	0.0094	0.0069
Minor appliances	0.0001	0.0020	0.0013	0.0336	0.0212	0.0001
Electronics	0.0005	0.0067	0.0005	0.0176	0.0091	0.0854
Vehicle purchases	0.0023	0.0332	0.0004	0.9844	0.0092	0.0008
<b>Nondurables</b>	<b>0.0027</b>	<b>0.0000</b>				
Food, alcohol, and tobacco	0.0009	0.0007	0.0009	0.0008	0.0001	0.5506
Food	0.0007	0.0035	0.0007	0.0063	0.0011	0.1427
Alcohol	0.0000	0.5658	0.0000	0.8426	-0.0052	0.4628
Tobacco	0.0001	0.1492	0.0000	0.7804	0.0083	0.0847
Food away from home	0.0001	0.1018	0.0001	0.5328	0.0138	0.0136
Apparel	-0.0002	0.3891	-0.0004	0.2278	0.0140	0.0763
Trips	0.0008	0.0000	0.0020	0.1402	0.0243	0.0024
Transportation	0.0011	0.0003	0.0011	0.0006	0.0022	0.2173
Gasoline	0.0002	0.0427	0.0002	0.0358	0.0007	0.7920
Other vehicle expenses	0.0008	0.0047	0.0009	0.0187	0.0122	0.0918
Public transportation	0.0001	0.0177	0.0004	0.2704	0.0135	0.0061
Entertainment	0.0000	0.8229	0.0000	0.8136	-0.0024	0.5228
Fees, admissions, toys, and sports	0.0000	0.9810	-0.0001	0.7525	0.0057	0.4336
Personal care services	0.0000	0.1728	0.0001	0.1741	0.0052	0.4177
Reading	0.0000	0.9818	0.0000	0.5927	0.0063	0.4132
Other nondurables	0.0001	0.6995	-0.0001	0.7713	0.0068	0.3439
Child support, alimony, and charity	0.0002	0.1614	0.0002	0.6355	0.0143	0.0436
Pensions, insurance, and social security	-0.0001	0.1145	-0.0004	0.4439	-0.0154	0.0120
Children's clothing only among families with children	0.0002	0.1490	-0.0007	0.0197	0.0415	0.0000

Notes: EITC means earned income tax credit. Children's clothing is a portion of the apparel subcategory.

Source: Authors' calculations based on data from the U.S. Bureau of Labor Statistics, *Consumer Expenditure Survey*.

than in September, EITC families spend about the same in February and September. We do not investigate conditional or discrete expenditure because the probability of making nondurable goods expenditure is nearly 1 in a given month.

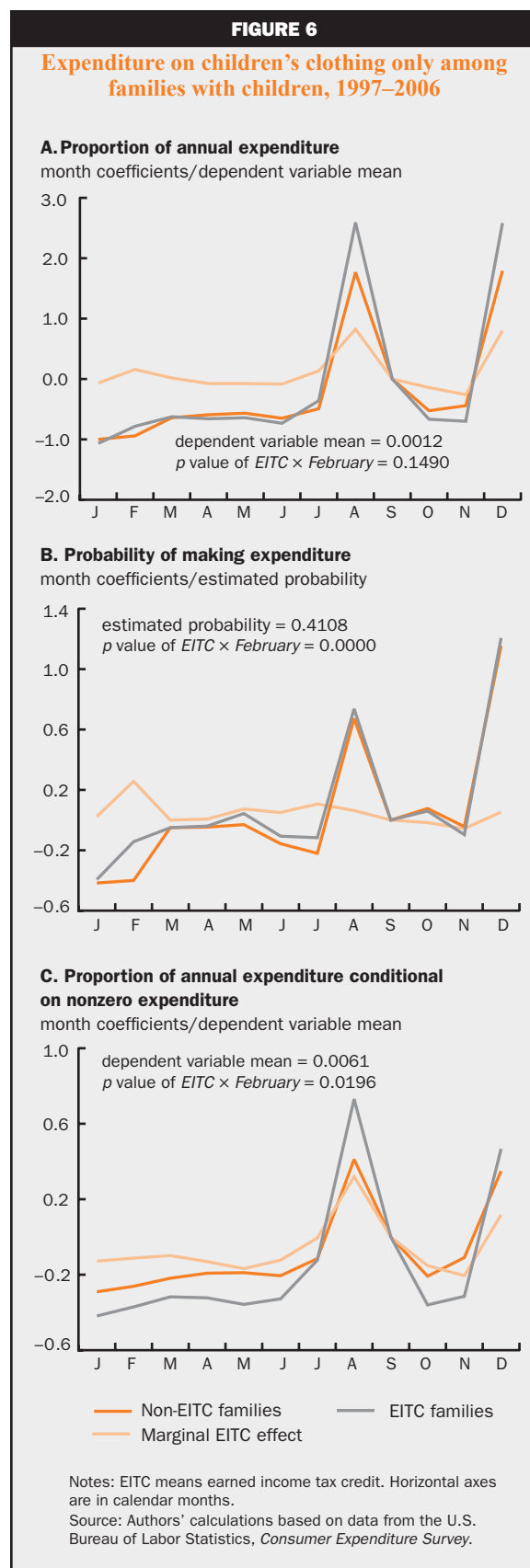
In figure 6, we present results for a subset of nondurables that is particularly relevant to the EITC's goals: expenditures on children's clothes. We estimate these models only for families with children so that the non-EITC control group is not dominated by childless families. Overall seasonal patterns between EITC families with children and non-EITC families with children are very similar, exhibiting a large increase in expenditures before school starts in September and during the holiday season (panel A). The EITC families are more likely to buy children's clothes in February than non-EITC families (panel B), but since they spend a slightly lower proportion of their total annual expenditure conditional on buying children's clothes (panel C), we do not find a statistically significant unconditional effect.

In figure 7, we present results for the nondurables portion of transportation. This includes gasoline, local public transportation, and car expenses outside of vehicle purchases. We find that EITC eligible households spend about 4 percent more in February than September, while noneligible households spend about 3 percent less (panel A). Most of this difference arises from higher spending conditional on positive expenditure (panel C). If we look at the first column of table 4, we find that transportation spending increases in February are the largest single contributor to the overall nondurables increase. From table 4, we also observe that EITC households spend relatively more on food and on trips than noneligible households in February.

### Durable goods

Figure 8 presents results for all durable goods. The difference in expenditure patterns between EITC and non-EITC families in February is much more pronounced than for nondurable goods. While non-EITC families spend about 8 percent less on durables in February than in September, EITC families spend about 18 percent more (panel A). The EITC families are significantly more likely both to make a durable goods purchase in February and to spend more conditional on making a purchase (panels B and C, respectively).

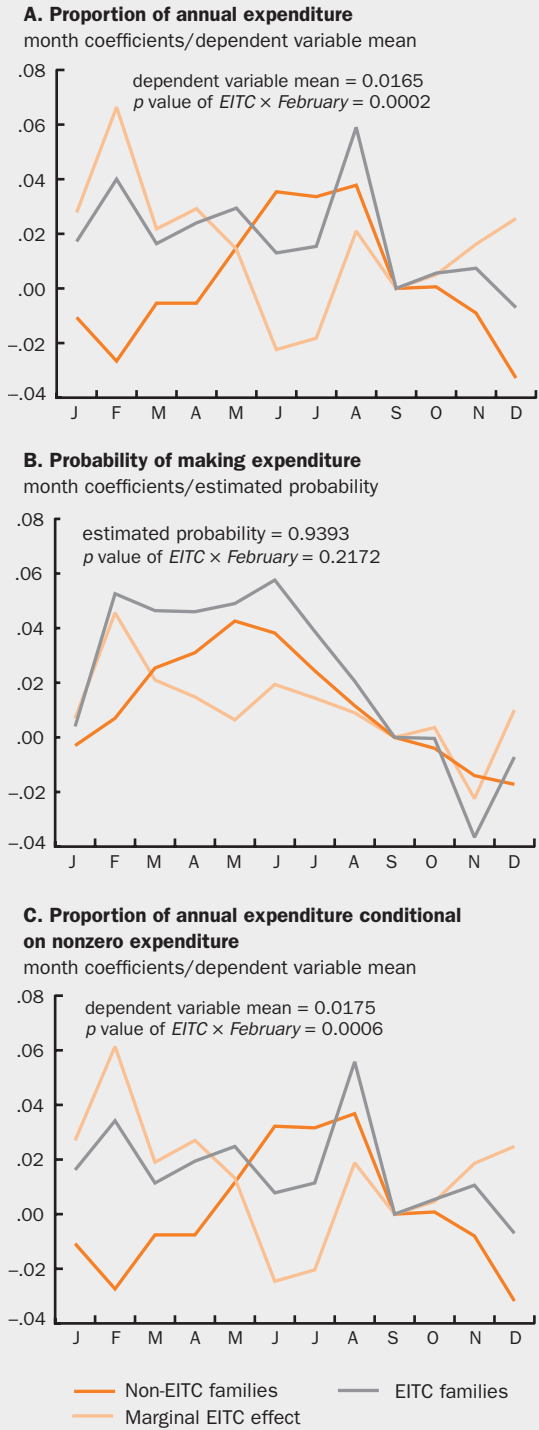
We now examine the subcategories of durable goods that drive the patterns depicted in figure 8. Figure 9 presents results for new and used vehicle purchases.<sup>15</sup> While non-EITC families spend about 17 percent less on vehicles in February than in September, EITC families spend 18 percent more (panel A), for a statistically significant difference of 35 percent ( $p$  value = 0.0332). This difference is entirely attributable to the fact that





**FIGURE 7**

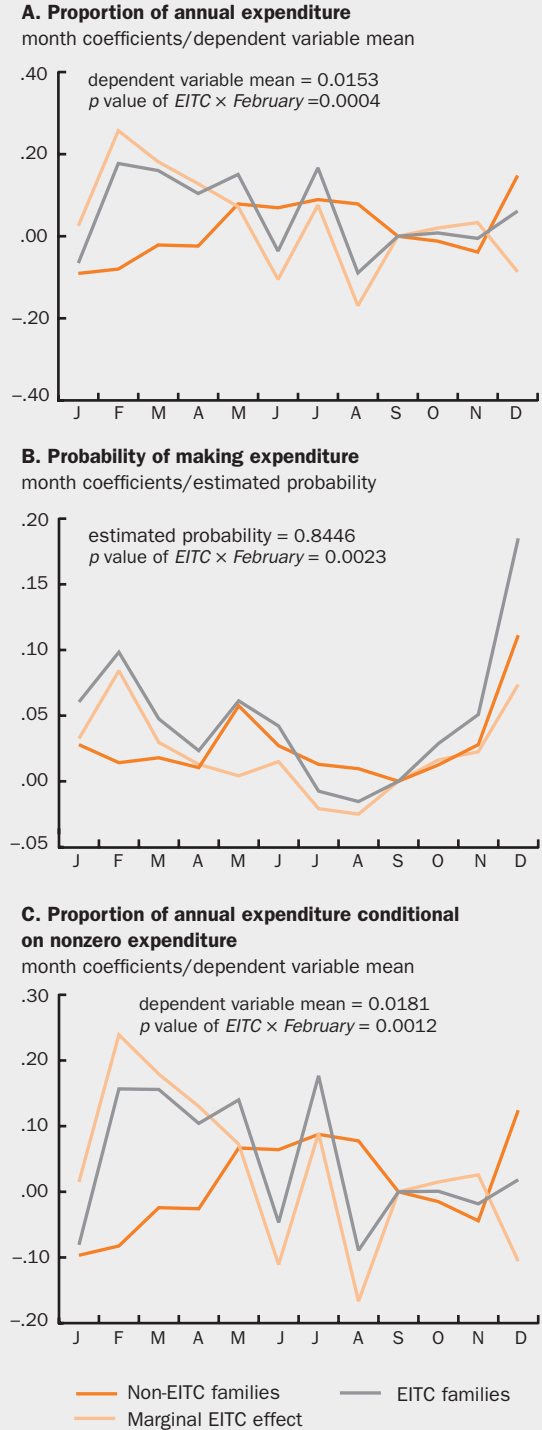
**Expenditure on transportation excluding vehicle purchases, 1997–2006**



Notes: EITC means earned income tax credit. Horizontal axes are in calendar months.  
Source: Authors' calculations based on data from the U.S. Bureau of Labor Statistics, *Consumer Expenditure Survey*.

**FIGURE 8**

**Durable goods expenditures, 1997–2006**

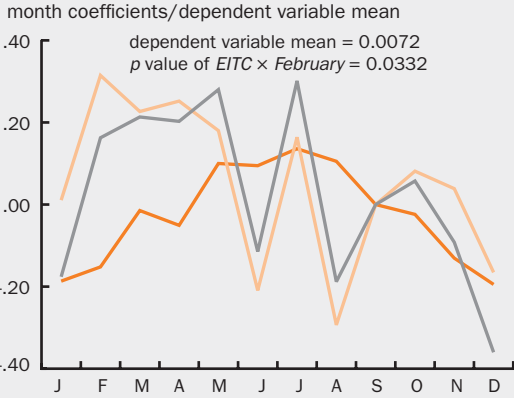


Notes: EITC means earned income tax credit. Horizontal axes are in calendar months.  
Source: Authors' calculations based on data from the U.S. Bureau of Labor Statistics, *Consumer Expenditure Survey*.

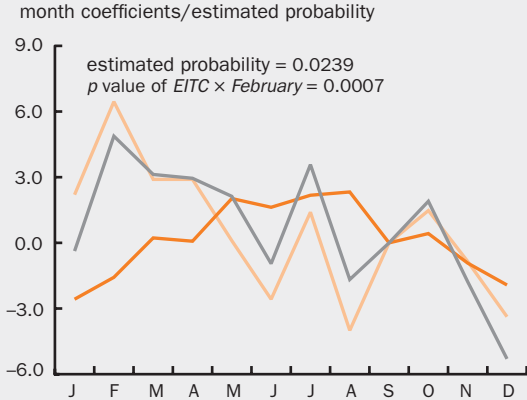
**FIGURE 9**

**Vehicle purchases expenditures, 1997–2006**

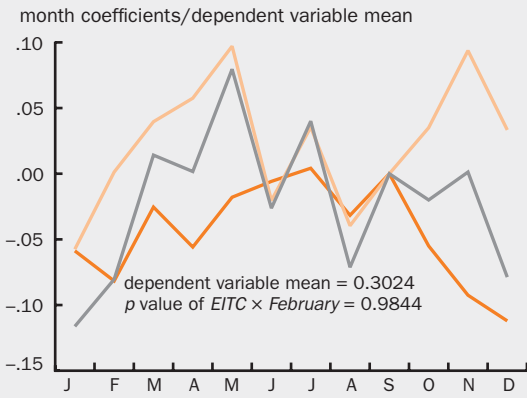
**A. Proportion of annual expenditure**



**B. Probability of making expenditure**



**C. Proportion of annual expenditure conditional on nonzero expenditure**



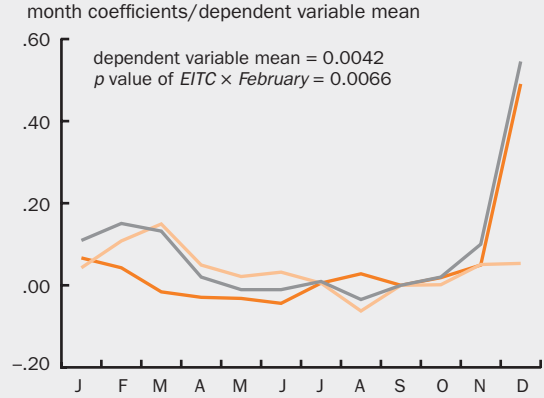
— Non-EITC families    — EITC families  
— Marginal EITC effect

Notes: EITC means earned income tax credit. Horizontal axes are in calendar months.  
Source: Authors' calculations based on data from the U.S. Bureau of Labor Statistics, *Consumer Expenditure Survey*.

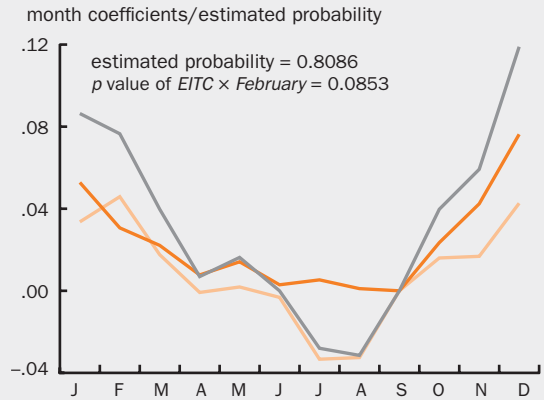
**FIGURE 10**

**Consumer electronics expenditures, 1997–2006**

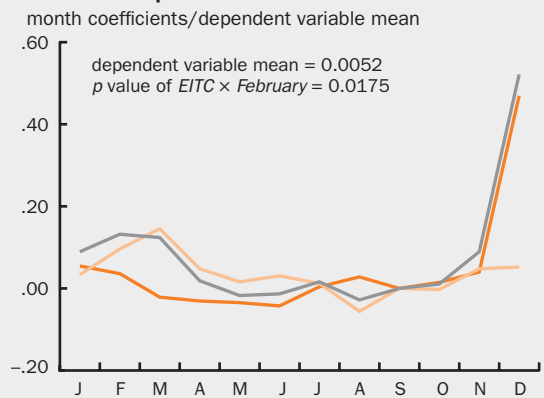
**A. Proportion of annual expenditure**



**B. Probability of making expenditure**



**C. Proportion of annual expenditure conditional on nonzero expenditure**



— Non-EITC families    — EITC families  
— Marginal EITC effect

Notes: EITC means earned income tax credit. Horizontal axes are in calendar months.  
Source: Authors' calculations based on data from the U.S. Bureau of Labor Statistics, *Consumer Expenditure Survey*.

relative to September, EITC families are more than 600 percent more likely than non-EITC families to buy a car in February (panel B). This difference is about twice as large in February as in either January or March. These findings are also consistent with the research of Adams, Einav, and Levin (2007), which shows high demand for subprime auto loans in tax rebate season among households likely to be receiving an EITC refund. Interestingly, though, among families making a vehicle purchase in February (panel C), all families spend the same proportion of their annual expenditure on these goods ( $p$  value = 0.9844). Recall that in dollars, this amount is considerably smaller for EITC families.

Figure 10 graphs the coefficients from models of spending on consumer electronics, which include television sets, computers, and video and music players. Considering all observations, non-EITC households spend about 5 percent more on consumer electronics in February than in September, and EITC households spend about 15 percent more (panel A). However, the February effect is relatively small compared with the overall February effect for durable goods and substantially smaller than the effect for vehicles.

Results for other subcategories of durable goods are similar to the findings for electronics. In February, EITC eligible households spend more than noneligible

households on both household goods and appliances, but the magnitude of these effects is smaller than the magnitude of the effect for vehicles.

## Conclusion

The results presented here indicate that EITC families spend at least a portion of their refund immediately upon receipt. Consistent with Barrow and McGranahan (2000), we find that recipients spend more on durables than on nondurables in response to the EITC. In particular, recipients are far more likely to purchase vehicles after receiving EITC refunds. The EITC increases relative average monthly spending on vehicles in February by about 35 percent for EITC families compared with their non-EITC counterparts. Within nondurables, expenditure increases are concentrated in transportation. Given the crucial role of access to transportation in promoting work, this leads to the conclusion that recipient spending patterns support the program's prowork goals. The EITC recipients are also more likely to spend money within the other durable goods categories, as well as on trips and food.

In future work, we hope to further analyze the consumption effects of the EITC by taking advantage of differences in state EITCs and by exploiting expansions in the EITC since its inception, as well as the changes in the timing of EITC payments.

## NOTES

<sup>1</sup>The 2005 CES contains data for the first quarter of 2006.

<sup>2</sup>A qualifying child must meet three requirements. First, this individual must be a child, stepchild, foster child, sibling, half sibling, stepsibling, or a descendent of a sibling of the tax filer. Second, the qualifying child must be younger than 19 at the end of the year, younger than 24 and a full-time student, or permanently disabled. Third, the qualifying child must live with the tax filer in the U.S. for at least six months out of the year. If two tax filers can claim one qualifying child, they can choose which one claims the child, but they both cannot claim the same child (U.S. Internal Revenue Service, 2006a). Starting in 2002, some married taxpayers filing jointly had higher benefits than singles with the same income and number of children.

<sup>3</sup>For e-filing, the e-file window needs to be open. This occurs in early January and happened on January 12, 2007.

<sup>4</sup>This was determined from authors' calculations based on data from the U.S. Internal Revenue Service (2006b).

<sup>5</sup>These figures for tax year 2004 are based on calculations using U.S. Internal Revenue Service (2006b) data. We assume that all overpayment refunds not due to the EITC are given to non-EITC recipients. The 26 percent of non-EITC taxpayers who did not receive a refund are included as zeros in this calculation.

<sup>6</sup>U.S. Social Security Administration (2006); and U.S. Department of Agriculture, Food and Nutrition Service (2008).

<sup>7</sup>The nondurables portion of transportation consists of gasoline and motor oil (42 percent), other vehicle expenses (49 percent), and public transportation (9 percent), according to the U.S. Bureau of Labor Statistics (2007).

<sup>8</sup>A consumer unit is defined to be an individual or a group of individuals who are either related or use their income to make joint expenditures on two of three categories—housing, food, or other living expenses.

<sup>9</sup>In future work, we plan to take advantage of changes in the timing of EITC payments and of expansions in the EITC to further investigate consumption responses to the program.

<sup>10</sup>In 2004, the credit for childless families accounted for only 3 percent of EITC payments despite representing 21 percent of returns receiving the EITC (U.S. Internal Revenue Service, 2006b).

<sup>11</sup>Our method of dealing with qualifying children could falsely impute EITC eligibility or inflate refund amounts for CUs with children and multiple, unrelated TUs. This is only a potential problem for the 4 percent of CUs that contain multiple TUs, have any qualifying children, and were assigned the EITC. Furthermore, if EITC eligibility “truly” has an impact on expenditure, then misallocating households into the EITC group should bias our results away from finding a difference in expenditure seasonality between eligible and noneligible CUs.

<sup>12</sup>Throughout the analysis, we rely on the monthly data in the CES. In some cases the monthly information is unreliable because of the random attribution of some expenditure to months in the survey. This attribution would likely operate in the same manner for EITC recipient and nonrecipient households.

<sup>13</sup>For households with 12 observations,  $X_{i,Annual} = \sum_{t=1}^{12} X_{i,t}^{Total}$ . In order to adjust monthly expenditure proportions for households with fewer than 12 observations, we regress  $X_{it}^{Total}$  on household characteristics for 12-month households only and then generate predicted expenditure proportions for all households. The sum of these predicted monthly proportions gives the expected proportion of annual expenditures that we actually observe for households with fewer than 12 observations. Thus, we estimate true annual expenditures by dividing the sum of  $m$  ( $m < 12$ ) observed expenditures by the

sum of  $m$  expected monthly proportions:  $\frac{\sum_{t=1}^m X_{it}^{Total}}{\sum_{t=1}^m E\left(\frac{X_{it}^{Total}}{X_{i,Annual}}\right)}$ . We use this

expression as the denominator of monthly expenditure proportions for households with fewer than 12 observations. It is because of this adjustment that average monthly expenditures are not equal to 1/12, or 8.33 percent, in table 2. We do not adjust the estimated standard errors in our regressions for this imputation.

<sup>14</sup>In their study of retail markdowns in Ann Arbor, Michigan, Warner and Barsky (1995) note that “prices are indeed lowest in January, but tend to return in February to December’s level.” We do not correct for the fact that February has fewer days than other months, which should, all else being equal, reduce February expenditures for both EITC recipient and nonrecipient households.

<sup>15</sup>According to the CES documentation, vehicle expenditures are defined as the purchase price minus the trade-in value on new and used domestic and imported cars and trucks and other vehicles, including motorcycles and private planes.

## REFERENCES

- Adams, William, Liran Einav, and Jonathan D. Levin**, 2007, "Liquidity constraints and imperfect information in subprime lending," National Bureau of Economic Research, working paper, No. 13067, April.
- Barrow, Lisa, and Leslie McGranahan**, 2000, "The effects of the earned income credit on the seasonality of household expenditures," *National Tax Journal*, Vol. 53, No. 4, part 2, December, 1211–1244.
- Blumenberg, Evelyn, and Margy Waller**, 2003, "The long journey to work: A federal transportation policy for working families," Brookings Institution Series on Transportation Reform, Brookings Institution, Center on Urban and Metropolitan Policy, report, July.
- Frost, Jonas Martin, III (D-TX)**, 1993, speaking before the U.S. House of Representatives, 103rd Cong., 1st sess., *Congressional Record*, Vol. 139, July 29, p. H5502.
- Grogger, Jeffrey**, 2003, "The effects of time limits, the EITC, and other policy changes on welfare use, work, and income among female-headed families," *Review of Economics and Statistics*, Vol. 85, No. 2, pp. 394–408.
- Eissa, Nada, and Jeffrey B. Liebman**, 1996, "Labor supply response to the earned income tax credit," *Quarterly Journal of Economics*, Vol. 111, No. 2, May, pp. 605–637.
- Linnenbrink, Mary, Michael Rupured, Teresa Mauldin, and Joan Koonce Moss**, 2006, "The earned income tax credit: Experiences from and implications of the voluntary income tax assistance program in Georgia," in *2006 Eastern Family Economics and Resource Management Association Conference Proceedings*, section A, pp. 11–16, available at <http://mrupured.myweb.uga.edu/conf/2.pdf>.
- Mammen, Sheila, and Frances Lawrence**, 2006, "Use of the earned income tax credit by rural working families," in *2006 Eastern Family Economics and Resource Management Association Conference Proceedings*, section B, pp. 29–37, available at <http://mrupured.myweb.uga.edu/conf/4.pdf>.
- Meyer, Bruce D., and Dan T. Rosenbaum**, 2001, "Welfare, the earned income tax credit, and the labor supply of single mothers," *Quarterly Journal of Economics*, Vol. 116, No. 3, August, pp. 1063–1114.
- Smeeding, Timothy M., Katherin Ross Phillips, and Michael O'Connor**, 2000, "The EITC: Expectation, knowledge, use, and economic and social mobility," *National Tax Journal*, Vol. 53, No. 4, part 2, December, pp. 1187–1210.
- U.S. Bureau of Labor Statistics**, 2007, "Consumer expenditures in 2005," report, Washington, DC, No. 998, February, available at [www.bls.gov/cex/csxann05.pdf](http://www.bls.gov/cex/csxann05.pdf), accessed on February 5, 2008.
- U.S. Department of Agriculture, Food and Nutrition Service**, 2008, "Food Stamp Program," report, Alexandria, VA, February 19, available at [www.fns.usda.gov/fsp/faqs.htm](http://www.fns.usda.gov/fsp/faqs.htm), accessed on March 3, 2008.
- U.S. Government Accounting Office**, 2001, "Earned income tax credit participation," report, Washington, DC, No. GAO-02-290R, December 14.
- U.S. House of Representatives, Committee on Ways and Means**, 2004, *2004 Green Book: Background Material and Data on the Programs Within the Jurisdiction of the Committee on Ways and Means*, Washington, DC, April.
- U.S. Internal Revenue Service**, 2006a, "Publication 596 (2006), earned income credit (EIC)," report, Washington, DC, available at [www.irs.gov/publications/p596/index.html](http://www.irs.gov/publications/p596/index.html), accessed on August 6, 2007.
- \_\_\_\_\_, 2006b, "SOI tax stats—Individual income tax returns publication 1304 (complete report)," report, Washington, DC, available at [www.irs.gov/taxstats/indtaxstats/article/0,,id=134951,00.html](http://www.irs.gov/taxstats/indtaxstats/article/0,,id=134951,00.html), accessed on June 4, 2007.
- \_\_\_\_\_, 2006c, "SOI tax stats—Issuing refunds," report, Washington, DC, available at [www.irs.gov/taxstats/compliancestats/article/0,,id=97270,00.html](http://www.irs.gov/taxstats/compliancestats/article/0,,id=97270,00.html).
- U.S. Social Security Administration**, 2006, "Highlights and trends," in *Annual Statistical Supplement, 2005*, Baltimore, MD, February, pp. 1–8, available at [www.ssa.gov/policy/docs/statcomps/supplement/2005/highlights.pdf](http://www.ssa.gov/policy/docs/statcomps/supplement/2005/highlights.pdf), accessed on March 3, 2008.
- Warner, Elizabeth, and Robert B. Barsky**, 1995, "The timing and magnitude of retail store markdowns: Evidence from weekends and holidays," *Quarterly Journal of Economics*, Vol. 110, No. 2, May, pp. 321–352.



# Are inflation targets good inflation forecasts?

Marie Diron and Benoît Mojon

## Introduction and summary

The growing use of inflation targeting and other forms of quantified inflation objectives has marked the history of monetary policy since 1990. Indeed, a majority of industrialized countries have either adopted some form of inflation targeting or, most notably for the 15 countries that have adopted the euro, defined a quantified inflation objective. In the United States, the Federal Reserve System aims to conduct the nation's monetary policy by influencing the monetary and credit conditions in the economy in pursuit of "maximum employment, stable prices, and moderate long-term interest rates."<sup>1</sup> The Fed does not have an inflation target.

An inflation target is a numerical point or range for the inflation of a given price index that the central bank declares to be its objective for inflation. For instance, the Bank of Canada aims to keep inflation at the 2 percent target. And the European Central Bank (ECB) aims to keep inflation below but close to 2 percent. Central banks that have a quantified inflation objective do structure the communication of their monetary policy around this objective.<sup>2</sup> Table 1 shows how various central banks currently define their inflation objectives, as reported on the central banks' websites. Table 2 shows when these targets were adopted and how they have changed. Inflation point targets and the midpoints of inflation target ranges are usually between 2 percent and 2.5 percent. These targets were first introduced between the early 1990s and the early 2000s. There is a broad consensus among economists that, as shown in figure 1, countries that have adopted an inflation target have stabilized inflation close to the inflation target.

In theory, a major virtue of quantified inflation objectives is to anchor inflation expectations—a key ingredient for the success of monetary policy. Stabilizing inflation expectations is important<sup>3</sup> because prices and wages adjust relatively infrequently (for the most

up-to-date evidence, see Dhyne et al., 2005; Fabiani et al., 2005; Vermeulen et al., 2007; and the references therein). The people and institutions in the economy (we call these economic agents) usually set prices and wages over some horizon, and the level of these prices and wages would reflect their expectation of the evolution of inflation. If these economic agents know what the official inflation target is and the target is credible, they will expect the general price level to grow at the rate of the preannounced objective of the central bank. This expectation in itself then helps to deliver realized inflation close to the target.

While many economists find this argument to be convincing, there has been little research so far on whether the central banks' targets actually do a better job at forecasting inflation than other inflation benchmarks. In this article, we evaluate the potential benefits of inflation targets by comparing the performance of benchmark forecasts of inflation (model-based and published forecasts) and forecasts that are set equal to the inflation target. We conduct this comparison of forecast performance for the euro area, Australia, Canada, New Zealand, Norway, Sweden, Switzerland, and the United Kingdom, all of which have established inflation targets as shown in table 1.

*Marie Diron is an economist at Oxford Economics. She worked on this project while she was at the European Central Bank. Benoît Mojon is an economist at the Federal Reserve Bank of Chicago, and is on leave from the European Central Bank. The authors thank Gonzalo Camba-Mendez, Han Choi, Michael Ehrmann, Gabriel Fagan, Jonas Fisher, Alejandro Justiniano, Simone Manganelli, Sergio Nicoletti-Altimari, Athanasios Orphanides, Anna Paulson, Frank Smets, Lars Svensson, David Vestin, and participants in a Chicago Fed research seminar and the Eurosystem Inflation Persistence Network September 2005 meeting for comments and suggestions. The views expressed here are the authors' and do not necessarily reflect the views of the European Central Bank.*

**TABLE 1****Inflation objectives in selected Organization for Economic Cooperation and Development countries and in the euro area**

Euro area	The primary objective of the European Central Bank's (ECB) monetary policy is to maintain price stability. The ECB aims at (harmonized index of consumer prices, or HICP) inflation rates of below, but close to, 2 percent over the medium term.
Australia	In pursuing the goal of medium-term price stability, both the bank and the government agree on the objective of keeping consumer price inflation between 2 percent and 3 percent, on average, over the cycle. This formulation allows for the natural short-run variation in inflation over the business cycle while preserving a clearly identifiable performance benchmark over time.
Canada	The Bank of Canada aims to keep inflation at the 2 percent target, the midpoint of the 1 percent to 3 percent inflation-control target range. This target is expressed in terms of total Consumer Price Index (CPI) inflation, but the bank uses a measure of core inflation as an operational guide. Core inflation provides a better measure of the underlying trend of inflation and tends to be a better predictor of future changes in the total CPI.
New Zealand	The Reserve Bank uses monetary policy to maintain price stability as defined in the policy targets agreement (PTA). The current PTA requires the bank to keep inflation between 1 percent and 3 percent on average over the medium term. The bank implements monetary policy by setting the official cash rate (OCR), which is reviewed eight times a year.
Norway	The government has defined an inflation target for monetary policy in Norway. The operational target is an inflation rate of 2.5 percent over time (with annual consumer price inflation of approximately 2.5 percent over time).
Sweden	According to the Sveriges Riksbank Act, the objective of monetary policy is to "maintain price stability." The Riksbank [or the central bank of Sweden] has interpreted this objective to mean a low, stable rate of inflation. More precisely, the Riksbank's objective is to keep inflation around 2 percent per year, as measured by the annual change in the Consumer Price Index (CPI). There is a tolerance range of plus/minus 1 percentage point around this target. At the same time, the range is an expression of the Riksbank's ambition to limit such deviations. In order to keep inflation around 2 percent, the Riksbank adjusts its key interest rate, the repo rate.
Switzerland	The Swiss National Bank equates price stability with a rise in the national Consumer Price Index (CPI) of less than 2 percent per annum. In so doing, it takes account of the fact that not every price movement is necessarily inflationary in nature. Furthermore, it believes that inflation cannot be measured accurately. Measurement problems arise, for example, when the quality of goods and services improves. Such changes are not properly accounted for in the CPI; as a result, the measured level of inflation will tend to be slightly overstated.
United Kingdom	A principal objective of any central bank is to safeguard the value of the currency in terms of what it will purchase. Rising prices—inflation—reduces the value of money. ... In May 1997, the government gave the bank independence to set monetary policy by deciding the level of interest rates to meet the government's inflation target—currently 2 percent. [The inflation target of 2 percent is expressed in terms of an annual rate of inflation based on the Consumer Prices Index (CPI).]

Sources: European Central Bank, [www.ecb.int/mopo/html/index.en.html](http://www.ecb.int/mopo/html/index.en.html); Reserve Bank of Australia, [www.rba.gov.au/MonetaryPolicy/](http://www.rba.gov.au/MonetaryPolicy/); Bank of Canada, [www.bank-banque-canada.ca/en/monetary/monetary\\_main.html](http://www.bank-banque-canada.ca/en/monetary/monetary_main.html); Reserve Bank of New Zealand, [www.rbnz.govt.nz/monpol/index.html](http://www.rbnz.govt.nz/monpol/index.html); Norges Bank, [www.norges-bank.no/Pages/Section\\_\\_\\_11330.aspx](http://www.norges-bank.no/Pages/Section___11330.aspx); Sveriges Riksbank, [www.riksbank.com/templates/SectionStart.aspx?id=10602](http://www.riksbank.com/templates/SectionStart.aspx?id=10602); Swiss National Bank, [www.snb.ch/en/iabout/monpol](http://www.snb.ch/en/iabout/monpol); and Bank of England, [www.bankofengland.co.uk/monetarypolicy/index.htm](http://www.bankofengland.co.uk/monetarypolicy/index.htm).

TABLE 2

## Characteristics of the inflation quantified objectives

	First introduction	Range (%)	Point target (%)	Horizon	Most recent modifications	Range (%)	Point target (%)	Horizon	Target forecast (%)	Forecast evaluation period
Euro area	November 1998	0–2	none	medium run	2003		close to 2 from below		1.9	2001:Q1–2007:Q4
Australia	April 1993	2–3	none	business cycle					2.5	1995:Q1–2007:Q4
Canada	February 1991	1–3	2	multiyear	1995			indefinite	2	1995:Q1–2007:Q4
New Zealand	March 1990	3–5	none		2003	1–3		business cycle	2	1995:Q1–2007:Q4
Norway	March 2001	none	2.5						2.5	2003:Q1–2007:Q4
Sweden	January 1993	1–3	2						2	1995:Q1–2007:Q4
Switzerland	1999	0–2	none						1	2001:Q1–2007:Q4
UK	October 1992	1–4 for RPIX	none		2004 switch to CPI	1–3 for CPI	2		2.5 RPIX 2.0 CPI	1995:Q1–2003:Q4 2004:Q1–2007:Q4

Notes: The euro itself was introduced in January 1999. However, the monetary policy strategy of the European Central Bank was announced in November 1998. In the United Kingdom, CPI is the Consumer Prices Index, and RPIX is the Retail Prices Index excluding mortgage interest payments. The last two columns define the target forecasts and the sample of forecast evaluation.  
Sources: Roger and Stone (2005); and Swiss National Bank, [www.snb.ch/en/iabout/monpol](http://www.snb.ch/en/iabout/monpol).

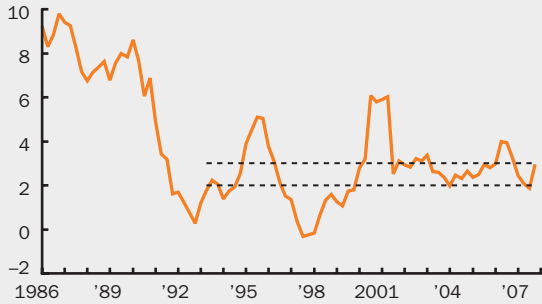
We also report results for the U.S., where inflation is often measured with the core Personal Consumption Expenditures (PCE) Price Index—a broad measure of consumer prices that excludes the more volatile and seasonal food and energy prices. Although the Federal Reserve does not have an inflation target, many market participants and economists assume that the U.S. central bank’s price stability mandate can be associated with numerical values for the core PCE inflation rate: Some have argued that this rate is close to 2 percent,<sup>4</sup> while others think that the Federal Reserve may have a “comfort zone” that is between 1 percent and 2 percent. Figure 2 shows that core PCE inflation was indeed close to these numerical values over the last decade. So, for comparison, we also assess the forecasting performance of two selected “constant forecast benchmarks” for the U.S.—one of core PCE inflation at 2 percent and the other at 1.5 percent (which is the midpoint of the alleged “comfort zone”).

Our results provide support for inflation targeting as a monetary policy strategy. In all the countries in our sample and in the euro area, forecasting that inflation will be at the inflation “target” implies a smaller forecasting error than alternative models. This is true for both one- and two-year horizon forecasts. Forecasting inflation to be at the target also beats the mean of professional economists’ forecasts published in *Consensus Forecasts* for the euro area, Canada, and Sweden, as well as for two-years-ahead forecasts in Switzerland and in the United Kingdom.<sup>5</sup> In the case of the U.S., forecasting core PCE inflation to be a constant benchmark (either at 2 percent or 1.5 percent) also implies a relatively small error on average over the past 12 years.

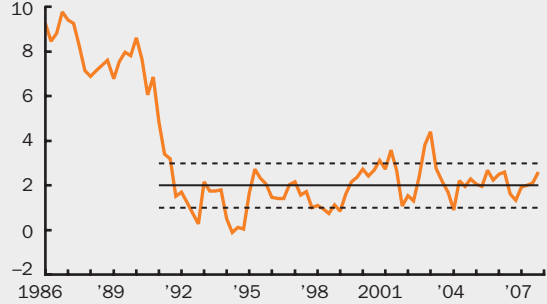
**FIGURE 1**

**Inflation and quantified inflation objectives**

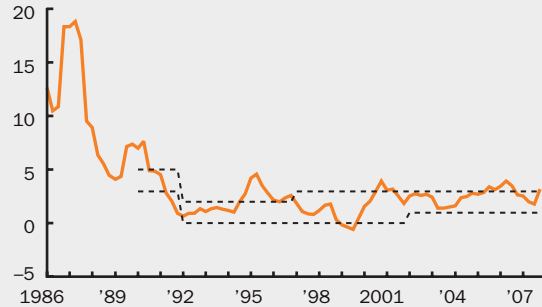
**A. Australia**  
percent



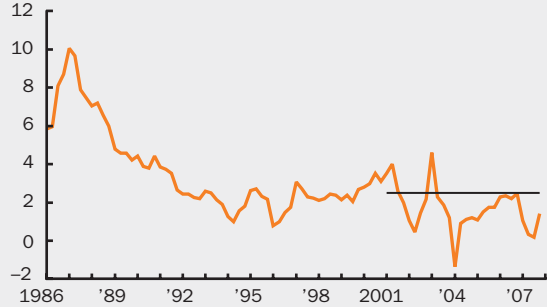
**B. Canada**  
percent



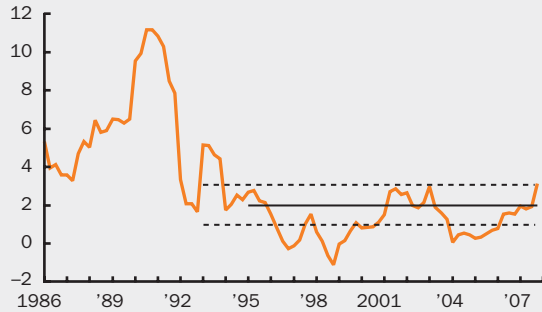
**C. New Zealand**  
percent



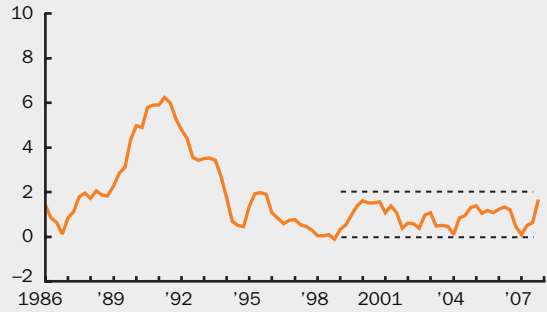
**D. Norway**  
percent



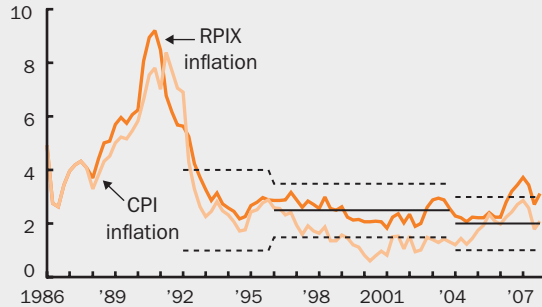
**E. Sweden**  
percent



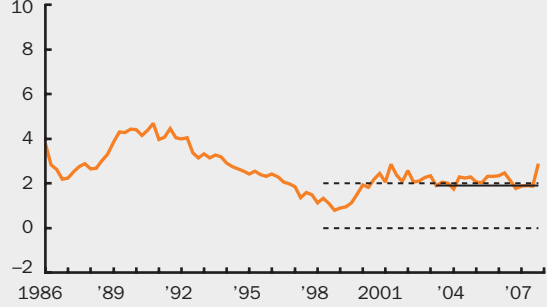
**F. Switzerland**  
percent



**G. United Kingdom**  
percent



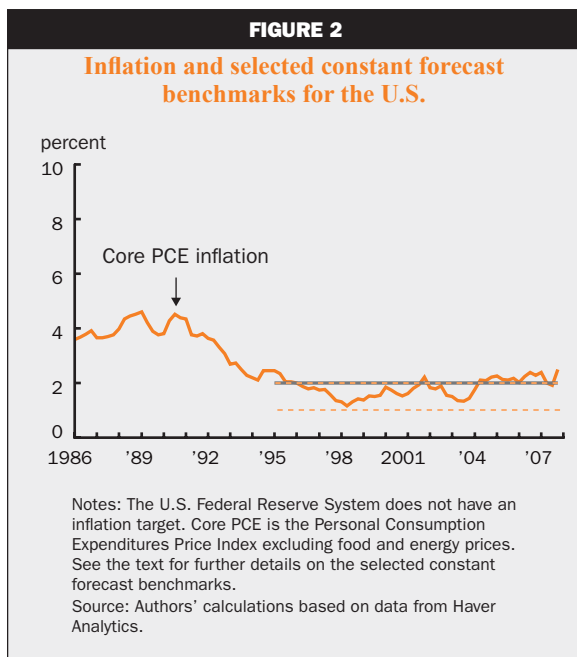
**H. Euro area**  
percent



----- Target ranges      ——— Point target

Notes: For panel G, in the United Kingdom, CPI is the Consumer Prices Index, and RPIX is the Retail Prices Index excluding mortgage interest payments. See tables 1 and 2 for details on the inflation objectives over the time period.

Sources: Roger and Stone (2005) and authors' calculations based on data from Haver Analytics.



To our knowledge, this article is the first one to show that, while inflation is never exactly at the target, the central bank's target has provided an *ex ante* reliable and, to a large extent, unbeatable inflation forecasting device in countries that have adopted a quantified inflation objective. When agents in the economy choose the inflation target as their expectation of future inflation, it is more likely that the target is actually hit or at least that low and stable inflation is maintained.

In the next section, we discuss the role of inflation targets in the formation of inflation expectations. Then, we describe the forecasting models and report the results of a "horse race" of inflation forecasts, comparing the error incurred by taking the target as a forecast with other widely used forecasting approaches.

### Rule-of-thumb expectations and inflation targets

The formation of inflation expectations plays a large role in the success of monetary policy. Since all prices and wages cannot be readjusted constantly, anchoring inflation expectations at a low level is essential to ensure price stability.

The academic debate on inflation expectations has centered on the operational mode of expectation formation. However, inflation expectations are not observable. As a result, several views on expectation formation that are mutually exclusive cannot easily be proven to be inconsistent with the data (Lindé, 2001).

The most popular view has long been that inflation expectations are rational. Rational expectations take two complementary meanings. First, expectations need to fulfill certain criteria to be *rational*. Thus, rational expectations cannot be systematically or persistently *wrong*. As a result, a good approximation of rational expectations is the result of a regression of future realizations of inflation on past and present observable economic variables. By construction, this procedure yields expectation errors that are zero on average. In addition, if the set of economic variables taken into account is comprehensive enough, this procedure is consistent with the requirement that expectations take into account all available information. The second meaning of rational expectations formulates that in any given model of the economy, agents form their expectations in a way that is consistent with the functioning of the model.

Although the assumption of rational expectations is frequently used in model construction and simulations, the empirical relevance is still controversial.<sup>6</sup> In particular, inflation expectations seem to depend significantly on past and present values of inflation (for example, Estrella and Fuhrer, 1999). Hence, some economists have advocated that expectations should be approximated by simpler expectation mechanisms, such as projecting inflation to be at the level observed in the past.

Note that such "rule-of-thumb" expectations are not necessarily irrational to the extent that rules deriving future inflation from its past values may be the most efficient use of current available information to derive the outlook for inflation. A good rationale for such a rule of thumb is precisely that inflation proves extremely difficult to forecast with multivariate economic models.<sup>7</sup> Simple rules of thumb may therefore optimally solve the trade-off between accuracy of the expectations and effort spent to derive them.<sup>8</sup> However, especially at times of persistent changes in inflation, such backward-looking rules will lead to recurring forecast errors of persistent signs.

In countries where the central bank has announced an inflation target, a natural rule of thumb consists of expecting that future inflation would be at the target. The forecast error of this rule of thumb is given by the deviation of realized inflation from the preannounced target. It is different from zero because the central bank cannot deliver an inflation rate that is exactly on target every period. However, the degree of forecast error will depend on which benchmarks are used and, in particular, on whether alternative forecasts are better or worse.



## How well do you forecast inflation if you believe in the central bank's target?

We first check how accurate “forecasts” of agents taking the central bank’s target for granted (henceforth, “target forecasts”) would perform compared with forecasts based on six alternative benchmarks: random walk; a track record or past mean inflation; three specifications of an autoregressive (AR) model of inflation, that is, a model where past and current inflation help forecast future inflation; and, finally, the mean inflation forecast published in *Consensus Forecasts*. These models, which are standard benchmarks in the forecast evaluation literature, have proved difficult to beat when trying to forecast inflation (Stock and Watson, 2003; and Banerjee, Marcellino, and Masten, 2003).

### *The quantified inflation objectives of central banks*

An inflation target takes the form of either a numerical value or a range for inflation and a commitment by the central bank to stabilize inflation close to the target level. Central banks that have a quantified inflation objective put it at the core of the communication of their monetary policy.<sup>9</sup> Table 1 (p. 34) reports the current (as of January 2008) definitions of the central banks’ inflation objectives taken from their websites. Table 2 (p. 35) shows the timing of the adoption of the targets and how they have changed over time. Figure 1 (p. 36) shows how the targets compare with actual inflation. The central banks’ inflation targets are now typically between 1 percent and 3 percent. Some central banks target a range (Australia, euro area, and Switzerland) and others a specific rate (Norway). Some banks have changed the definition of their objective over time (euro area and UK), while some have not (Australia). Changes have involved the range target (New Zealand and euro area) or even a change in the index for which the target is defined (UK).<sup>10</sup>

Going from the definition of the inflation targets to a target forecast requires two main assumptions. The first one is to choose a numerical value for the target. We choose the effective point target when the central bank has defined one (Canada, Norway, Sweden, and the UK from 1996 onward) or, in the case of countries with inflation range targets (Australia, New Zealand, Switzerland, and UK before 1996), we use the midpoint of the range in order to have a point estimate to which actual inflation can be compared (following Castelnuovo, Nicoletti-Altimari, and Rodríguez-Palenzuela, 2003). In the case of the euro area, the choice of a specific number for the inflation quantified objective is somewhat more delicate. In 1998, the ECB had defined its inflation objective as a positive inflation rate less than 2 percent over the

medium run. In May 2003, the ECB clarified its inflation objective as below but close to 2 percent.<sup>11</sup> We set the inflation objective for the euro area at 1.9 percent. While this choice is somewhat arbitrary and not necessarily in line with the perception of the ECB objective between 1999 and 2003, we believe it is consistent with the ECB strategy both before and after May 2003.

Finally, we also analyze the case of the U.S. As noted earlier, in contrast with the other central banks we study in this article, the Federal Reserve does not set a target for inflation. However, some observers have suggested that the Federal Reserve has an implicit target of 2 percent for core PCE inflation.<sup>12</sup> Some others consider that the Federal Reserve has a “comfort zone” that is between 1 percent and 2 percent. We thought it would be interesting to apply the same type of test to the forecasting performance of these working assumptions as we do to the official inflation targets of other countries, purely as an academic exercise. We therefore assess the size of the errors implied by forecasting core PCE inflation rates to be constant, either at 2 percent or 1.5 percent.

The second assumption we need to make is our choice of forecast evaluation period. Given the medium-term nature of the central banks’ objectives, which we interpret as a two-year horizon, we start our forecast evaluation period two years after the inflation target has been announced. Hence, in the case of Australia, where the inflation targeting strategy was launched in 1993, the forecast evaluation commences for forecasts of inflation for the first quarter of 1995. In the case of the euro area, we record forecast performances from 2001 onward. The level of the inflation forecast and the first date of the forecast evaluation are reported in table 2 (p. 35). In the case of the U.S., we arbitrarily start the forecasting evaluation in 1995.

### *Forecasting models*

The target forecast model (“Target” in tables 3–5 on pp. 41–42) is simply:

$$\pi^{t+h}_t = \pi^*,$$

where  $\pi_t = \left( \frac{P_t - P_{t-4}}{P_{t-4}} \right) \times 100$ , that is, it is the inflation rate for four quarters,  $h$  is either four quarters or eight quarters,  $P$  is the level of the price index, and  $\pi^*$  is the inflation quantified objective defined in the next to last column of table 2. The range of  $t + h$  dates for which the model is evaluated is given in the last column of table 2.

We should stress that the forecasts are the same whatever the horizon of the forecast. In this article, we report results only for  $h$  equal to four and eight quarters ahead.<sup>13</sup>

We compare the target forecast performance with the forecasts from our six alternative measures. The first of these is the random walk forecast; that is, we forecast inflation to be equal to the inflation observed over the year to the date when the forecast is made:

$$1) \quad \pi^{t+h}_t = \pi_t.$$

This forecasting model is sometimes formulated in the first difference of inflation, that is, changes in inflation from one period to the next. We stick to a level formulation, however, because inflation shows no trend for the sample over which the forecast evaluation is conducted. We also record the forecast performance of considering that future inflation would be well approximated by the average inflation level over the past five years (or 20 quarters). This naive forecast considers that the recent track record of inflation is the most informative about where inflation should be:

$$2) \quad \pi^{t+h}_t = \left( \sum_{i=1}^{20} \pi_{t-i} \right) / 20.$$

The main advantage with respect to the first model (equation 1) is that it may smooth out temporary noise in current inflation.

We then base inflation forecasts on three autoregressive models.<sup>14</sup> The first of these models simply relates current inflation to its lag levels, where the minimum lag is defined by the forecasting horizon. It is:

$$3) \quad \pi_t = C + \alpha\pi_{t-h} + \beta\pi_{t-h-4} + \varepsilon_t,$$

where  $C$ ,  $\alpha$ , and  $\beta$  are parameters and  $\varepsilon$  an error term, which are to be estimated recursively by ordinary least squares over the sample from the first quarter of 1986 to  $t$ .

This simplifies the forecasting procedure as it can be computed in one step rather than rolling the model over intermediate forecasts:

$$3a) \quad \pi^{t+h}_t = \hat{C} + \hat{\alpha}\pi_t + \hat{\beta}\pi_{t-4},$$

where the coefficients with  $\hat{\phantom{x}}$  have been estimated.

We also present results for two variants of this model. First, we formulate the autoregressive model on the first difference of inflation. This formulation

has the advantage that any change in the level of inflation would affect the forecasting performance of the model only for one observation (Banerjee, Marcellino, and Masten, 2003):

$$4) \quad \Delta\pi_t = C + \alpha\Delta\pi_{t-h} + \beta\Delta\pi_{t-h-1} + \varepsilon_t,$$

$$4a) \quad \pi^{t+h}_t = \pi_t + \hat{C} + \hat{\alpha}\Delta\pi_t + \hat{\beta}\Delta\pi_{t-1},$$

where  $\Delta\pi_t = \pi_t - \pi_{t-4}$ .

Second, in line with Labhard, Kapetanios, and Price (2007), we take into account potential breaks in the mean of inflation due to announcements of changes in the inflation objective by the central banks.<sup>15</sup> Hence, we enrich the AR model by allowing for changes in the intercept eight quarters after a change to the inflation targeting regime. In the case of Australia, for instance, the central bank announced its objective in 1993. We therefore include a one-step dummy taking a zero value before 1995 (1993 plus eight quarters) and one thereafter. We refer to this second set of models as “AR models with breaks.” They are:

$$5) \quad \pi_t = C + \sum C_i Ind_i + \alpha\pi_{t-h} + \beta\pi_{t-h-4} + \varepsilon_t,$$

$$5a) \quad \pi^{t+h}_t = \hat{C} + \sum \hat{C}_i Ind_i + \hat{\alpha}\pi_t + \hat{\beta}\pi_{t-4},$$

where  $Ind_i$  is a dummy variable that takes a value 1 from eight quarters after the announced change in the target.

We estimate the models from the first quarter of 1986 onward with year-on-year inflation rates.<sup>16</sup> The out-of-sample forecast evaluation is then carried out in *pseudo* real time. For example, the models are estimated from the first quarter of 1986 through the fourth quarter of 1994. Based on this estimation, we calculate forecasts at horizons four quarters and eight quarters ahead. Then we store the associated forecast errors and the one of taking the inflation forecast equal to the central bank’s quantified objective  $\pi^*$ , defined as follows:

$$\begin{aligned} \pi^{1995Q1}_{1994Q1} - \pi_{1995Q1} & \quad \text{and} \quad [\pi_{1995Q1} - \pi^*], \\ \pi^{1996Q1}_{1994Q1} - \pi_{1996Q1} & \quad \text{and} \quad [\pi_{1996Q1} - \pi^*]. \end{aligned}$$

The setup is brought forward sequentially by one quarter until the end of the evaluation sample.

Finally, we compare target forecasts to the *Consensus Forecasts* (hereafter, referred to as the

“consensus”), which is the mean of the forecasts surveyed by Consensus Economics Inc. from  $F$  professional forecasters.

$$6) \quad \pi^{t+h}_t = \left( \sum_{f=1}^F \pi^{t+h}_{t,f} \right) / F.$$

The consensus should represent informed forecasts produced on the basis of comprehensive information sets. Notably, respondents to the survey should be aware of the central bank’s inflation objective. In principle, differences between the views of economists on future inflation and the central bank’s stated objective can indicate that such an objective lacks credibility. However, inflation targets could be credible, albeit only in the medium run. For shorter horizons, economists may take into account a variety of factors that make actual inflation deviate temporarily from the target.

Data on the professionals’ forecasts for future inflation (for the current and following years) are available since 1990 for Canada, Norway, Sweden, Switzerland, and the UK and since 2002 for the euro area. However, we compile pre-2002 data as averages of country-level data (except Luxembourg), with fixed weights corresponding to the countries’ share in euro area consumption.<sup>17</sup> This current and following year framework differs from the rolling forecast horizon used to evaluate models 1–5. In order to compare the performance of the consensus with the degree of accuracy that target forecasts would have yielded had they been formed at the same time as the consensus surveys, we need to pay attention to the calendar of inflation data releases and the timetable of the consensus surveys. Publication delays of inflation data differ from one country to another and, in some cases, have changed over the period we study here. However, inflation data are typically published about one month after the end of the reference period. Meanwhile, the consensus survey results for a month,  $M$ , correspond to answers collected up to the middle of the previous month  $M - 1$ . We can therefore make the following comparisons. Consensus forecasts of inflation in the current year published in February rely on inflation data up to December of the previous year. Therefore, we need to forecast the whole year. We then compare these forecasts with four-quarters-ahead target forecasts. Similarly, we compare forecasts of inflation in the following year published in February with eight-quarters-ahead target forecasts.

## Results

Tables 3 and 4 show the mean absolute errors (MAEs) and the root mean square errors (RMSEs)<sup>18</sup>

of the target forecast and the five alternative quarterly models laid out in equations 1–5. Table 5 compares similar statistics for *Consensus Forecasts* and the target forecasts at an annual frequency. These statistics are computed for the forecast evaluation periods that begin either in 1995 or eight quarters after the instauration of the inflation quantified objective. For most countries, this is from 1995 through 2007—that is, for 52 quarterly forecasts for tables 3 and 4 and for 13 annual observations for table 5. However, the forecast evaluation starts only in 2001 for the euro area and Switzerland and in 2003 for Norway. In the case of the UK, the forecast evaluation is split in 2004 to reflect the change in the underlying price index.

For each row in tables 3–5, the numbers in bold indicate the smallest forecast errors. In tables 3 and 4, for each column we also compute the mean performance of each model across countries as the mean distance to the best performing model for each country.

Our results provide strong support for the inflation target forecasts as good devices for inflation forecasting. This is especially true at the eight quarters horizon, where forecasting the target systematically beats all other forecasting approaches (that is, has both the smallest MAE and smallest RMSE) except in the UK, where the best model for the Consumer Prices Index (CPI) is the simple AR model in equation 3. But one should take this particular result for the UK with a grain of salt because our evaluation is conducted only over 16 observations (from 2004 through 2007).

At the four quarters horizon, the performance of forecasting the target remains very impressive. This model is the best performing one in terms of either mean absolute errors (table 3) or root mean square errors (table 4) in Canada, Norway, and Switzerland. In both tables 3 and 4, the performance of forecasting the target is very close to the best model in most other cases: less than 0.05 percentage points above the best model in the euro area and Australia and less than 0.10 percentage points above the best model in New Zealand and Sweden. In the UK, the target forecast has an MAE and RMSE about 0.20 percentage points above the best model for the either the RPIX or the CPI. However, even at a four quarters horizon, the target forecast is the most robust approach in the sense that it is, on average, the closest to the best performing model of each country.

The target forecasts yield significantly more accurate forecasts than any of the autoregressive models and, hence, given the evidence reported in Stock and Watson (2003) and Banerjee, Marcellino, and Masten (2003), than most inflation forecast models (see note 7).

TABLE 3

## Mean absolute errors at four quarters and eight quarters horizons

	Four-quarters-ahead forecasts						Eight-quarters-ahead forecasts					
	Target	Alternative models					Target	Alternative models				
		1	2	3	4	5		1	2	3	4	5
Euro area	0.36	0.43	0.46	0.46	<b>0.33</b>	0.47	<b>0.36</b>	0.51	0.53	0.64	0.52	0.65
Australia	1.11	1.55	<b>1.10</b>	1.46	1.59	1.56	<b>1.11</b>	1.84	1.15	2.04	2.53	2.42
Canada	<b>0.65</b>	1.05	1.19	0.86	1.23	0.75	<b>0.65</b>	0.97	1.57	1.35	1.05	1.34
New Zealand	0.99	1.12	1.19	<b>0.91</b>	1.77	0.95	<b>0.99</b>	1.31	1.55	1.11	1.31	1.16
Norway	<b>0.91</b>	1.43	1.00	1.06	2.13	1.04	<b>0.91</b>	1.20	1.09	1.01	1.06	1.18
Sweden	1.06	<b>0.97</b>	1.70	1.48	1.29	1.21	<b>1.06</b>	1.42	2.10	2.70	1.41	1.87
Switzerland	<b>0.39</b>	0.60	0.42	0.61	0.54	0.55	<b>0.39</b>	0.71	0.45	1.17	0.70	0.84
UK CPI	0.57	0.45	0.69	<b>0.38</b>	0.46	0.46	0.57	0.63	0.77	<b>0.48</b>	0.82	0.66
UK RPIX	0.54	0.33	0.85	0.85	<b>0.32</b>	0.56	0.54	<b>0.41</b>	1.19	2.37	0.45	1.70
Mean difference with best model	<b>0.07</b>	0.22	0.29	0.23	0.41	0.18	<b>0.02</b>	0.30	0.45	0.73	0.39	0.61

Notes: In the United Kingdom, CPI is the Consumer Prices Index, and RPIX is the Retail Prices Index excluding mortgage interest payments. The forecast comparison is conducted in real time over the period 1995:Q1–2007:Q4 for Australia, Canada, New Zealand, and Sweden; over the period 2001:Q1–2007:Q4 for the euro area and Switzerland; over the period 2003:Q1–2007:Q4 for Norway; over the period 2004:Q1–2007:Q4 for UK CPI; and over the period 1995:Q1–2003:Q4 for UK RPIX. The numbers in bold indicate the best model.

Model 1 is the random walk, current year inflation; model 2 is the mean inflation over the last five years; model 3 is an autoregressive (AR) model in levels; model 4 is an AR model in first differences; and model 5 is an AR model in levels with breaks in the mean  $\pi$  inflation. See equations 1–5 in the text for the exact specification of the forecast.

TABLE 4

## Root mean square errors at four quarters and eight quarters horizons

	Four-quarters-ahead forecasts						Eight-quarters-ahead forecasts					
	Target	Alternative models					Target	Alternative models				
		1	2	3	4	5		1	2	3	4	5
Euro area	0.46	0.53	0.58	0.56	<b>0.42</b>	0.57	<b>0.46</b>	0.68	0.68	0.84	0.71	0.89
Australia	1.51	1.97	<b>1.50</b>	1.88	1.94	1.98	<b>1.51</b>	2.41	1.57	3.62	3.23	3.96
Canada	<b>0.86</b>	1.32	1.80	1.18	1.66	0.99	<b>0.86</b>	1.15	2.37	2.78	1.48	2.75
New Zealand	1.25	1.43	1.66	<b>1.18</b>	2.43	1.25	<b>1.25</b>	1.67	2.42	1.46	1.88	1.51
Norway	<b>1.23</b>	1.95	1.31	1.42	2.88	1.38	<b>1.23</b>	1.54	1.40	1.24	1.34	1.50
Sweden	1.28	<b>1.22</b>	2.19	2.00	1.55	1.59	<b>1.28</b>	1.61	2.74	3.78	1.72	2.99
Switzerland	<b>0.45</b>	0.72	0.49	0.73	0.64	0.66	<b>0.45</b>	0.83	0.56	1.44	0.81	1.11
UK CPI	0.70	0.53	0.82	<b>0.45</b>	0.54	0.53	0.70	0.73	0.92	<b>0.57</b>	0.95	0.77
UK RPIX	0.64	0.42	1.11	0.97	<b>0.41</b>	0.68	0.64	<b>0.49</b>	1.50	2.83	0.57	2.23
Mean difference with best model	<b>0.07</b>	0.26	0.42	0.29	0.53	0.21	<b>0.03</b>	0.34	0.68	1.16	0.51	1.07

Notes: In the United Kingdom, CPI is the Consumer Prices Index, and RPIX is the Retail Prices Index excluding mortgage interest payments. The forecast comparison is conducted in real time over the period 1995:Q1–2007:Q4 for Australia, Canada, New Zealand, and Sweden; over the period 2001:Q1–2007:Q4 for the euro area and Switzerland; over the period 2003:Q1–2007:Q4 for Norway; over the period 2004:Q1–2007:Q4 for UK CPI and over the period 1995:Q1–2003:Q4 for UK RPIX. The numbers in bold indicate the best model.

Model 1 is the random walk, current year inflation; model 2 is the mean inflation over the last five years; model 3 is an autoregressive (AR) model in levels; model 4 is an AR model in first differences; and model 5 is an AR model in levels with breaks in the mean  $\pi$  inflation. See equations 1–5 in the text for the exact specification of the forecast.

TABLE 5

Forecasting errors of target forecasts and *Consensus Forecasts*

	Mean absolute errors			Root mean square errors		
	Target	Consensus one-year-ahead forecasts	Consensus two-years-ahead forecasts	Target	Consensus one-year-ahead forecasts	Consensus two-years-ahead forecasts
Euro area	<b>0.27</b>	0.29	0.41	<b>0.31</b>	0.31	0.45
Canada	<b>0.29</b>	0.41	0.38	<b>0.36</b>	0.54	0.43
Sweden	<b>0.65</b>	0.69	0.67	<b>0.88</b>	0.95	0.88
Switzerland	0.24	<b>0.21</b>	0.41	<b>0.27</b>	0.33	0.48
UK	0.41	<b>0.34</b>	0.44	0.53	<b>0.42</b>	0.56

Notes: The forecast comparison is conducted in real time over the period 1995–2007 for Canada, Sweden, and the UK and over the period 2001–07 for the euro area and Switzerland. The consensus forecasts are the ones published in the February issue of *Consensus Forecasts* of the current year for one-year-ahead forecasts and the past year for the two-years-ahead forecasts. The numbers in bold indicate the best model.

TABLE 6

## Performance of selected constant forecast benchmarks and model-based forecasts of U.S. core PCE inflation

	Constant forecast benchmarks		Alternative models				
	1.5%	2.0%	1	2	3	4	5
<b>Mean absolute errors</b>							
Forecast horizon							
Four quarters	0.49	0.32	<b>0.30</b>	0.57	0.33	0.37	0.32
Eight quarters	0.49	<b>0.32</b>	0.38	0.74	0.65	0.34	0.64
<b>Root mean square errors</b>							
Forecast horizon							
Four quarters	0.40	0.38	<b>0.36</b>	0.70	0.40	0.47	0.37
Eight quarters	0.40	<b>0.38</b>	0.45	0.92	0.87	0.42	0.86

Notes: The U.S. Federal Reserve System does not have an inflation target. Core PCE is the Personal Consumption Expenditures Price Index excluding food and energy prices. The forecasting performance of the constant forecast benchmarks for U.S. core PCE inflation is purely illustrative. The forecast comparison is conducted in real time over the period 1995:Q1–2007:Q4. The numbers in bold indicate the best model.

Model 1 is the random walk, current year inflation; model 2 is the mean inflation over the last five years; model 3 is an autoregressive (AR) model in levels; model 4 is an AR model in first differences; and model 5 is an AR model in levels with breaks in the mean  $f$  inflation. See equations 1–5 in the text for the exact specification of the forecast.

Table 5 shows the MAEs and the RMSEs of target forecasts and the *Consensus Forecasts*, though this time using yearly observations. For two-years-ahead inflation forecasts, using the central bank's target has yielded smaller forecasting errors than the consensus forecasts in terms of either MAEs or RMSEs for all countries under review. This is also observed at one-year-ahead forecasts, except for the UK according to both the MAE and RMSE criteria and for Switzerland according to the MAE criterion.

One caveat applying to these results is that they are based on relatively short samples because of the availability of consensus forecasts for only the past 15 years and the even more recent switch to quantified inflation objectives by central banks. However, in our view, the paths of the forecasts obtained from the autoregressive models, the consensus, and the central

banks' targets suggest that the central banks' targets may constitute a new benchmark for forecast evaluation.

Finally, table 6 reports MAEs and RMSEs of the constant forecast benchmarks of 1.5 percent and 2 percent for U.S. core PCE inflation. Forecasting constant inflation at 2 percent has been the best at the eight quarters horizon and very close to the best at the four quarters horizon. These results show that, although the Federal Reserve does not have an inflation target, core PCE inflation has become remarkably stable in the U.S. since 1995.

Taking a broader perspective, our results provide concrete evidence of the success of preannounced quantified objectives for inflation. One possible interpretation of this success is that economic agents have indeed adopted the inflation target of the central bank as their inflation expectation for the general price level.



The inflation target may have become the focal point onto which decentralized inflation expectations have converged. This would occur if the target of the central bank is credible. That is, the central bank is always willing to take measures to ensure the target is reached over the specified horizon.

## Conclusion

We have shown that quantified inflation objectives can be used as rule-of-thumb forecasting devices. The experience of various countries that have adopted such objectives shows that, to a large extent, such a

rule of thumb yields smaller forecast errors than widely used forecasting models and the forecasts of professional experts published by Consensus Economics Inc. While inflation is never exactly at the target, the central banks' targets have provided ex ante reliable and, to a large extent, unbeatable inflation forecasting devices in countries that have adopted a quantified inflation objective. These findings suggest that the central banks that have set explicit targets for inflation have been successful in their often stated goal of anchoring inflation expectations.

---

## NOTES

<sup>1</sup>This is according to the Federal Reserve Act; see [www.federalreserve.gov/generalinfo/fract/sect02a.htm](http://www.federalreserve.gov/generalinfo/fract/sect02a.htm).

<sup>2</sup>See Roger and Stone (2005) for a detailed description of the inflation targeting in OECD (Organization for Economic Cooperation and Development) and emerging economies.

<sup>3</sup>See the discussion in Castelnuovo, Nicoletti-Altimari, and Rodriguez-Palenzuela (2003); Gürkaynak, Levin, and Swanson (2006); Levin, Natalucci, and Piger (2004); and Svensson (1999).

<sup>4</sup>A prominent example is Goodfriend (2007).

<sup>5</sup>*Consensus Forecasts*—a monthly publication by Consensus Economics Inc.—reports the forecasts of inflation by various investment banks and public and private organizations that have their own inflation forecasts. For further details, see [www.consensus-economics.com](http://www.consensus-economics.com).

<sup>6</sup>See, for instance, Rudd and Whelan (2006) and Sargent (1993).

<sup>7</sup>Stock and Watson (2003); Banerjee, Marcellino, and Masten (2003); and Banerjee and Marcellino (2003) show that multivariate models of inflation—that is, models where inflation dynamics are influenced by the evolution of other economic variables (output and unemployment)—hardly ever improve the forecast of inflation with respect to univariate nonstructural models of inflation. See also Fisher, Liu, and Zhou (2002) and Brave and Fisher (2004).

<sup>8</sup>The recent discussion of rational inattention (Sims, 2003; Mankiw and Reis, 2002; and Maćkowiak and Wiederholt, 2005) models explicitly how the cost of information processing could cause agents to restrict the information on which they base economic decisions.

<sup>9</sup>Again, see Roger and Stone (2005) for a detailed description of the inflation targeting in OECD and emerging economies.

<sup>10</sup>In December 2003, the UK's Chancellor of the Exchequer announced that the Bank of England would change its inflation target

from one based on the Retail Prices Index excluding mortgage interest payments (RPIX) to one based on the Consumer Prices Index (CPI)—also known there as the Harmonized Index of Consumer Prices (HICP).

<sup>11</sup>See Issing (2003).

<sup>12</sup>Goodfriend (2007).

<sup>13</sup>In a previous version of this article, we showed that the target forecast does not perform well at a one-quarter horizon—a result that is not surprising given that all central banks with an inflation target insist that inflation can be brought back to the target only over the medium run. In other words, it is widely agreed that monetary policy should not aim at cancelling the high frequency volatility of inflation.

<sup>14</sup>Other lag structures did not improve the forecasting results, so we use the simplest possible lag structure here.

<sup>15</sup>An obvious weakness of this model is that it assumes that the econometrician himself is convinced that the central bank announcement of a new target will immediately have an effect on the inflation process.

<sup>16</sup>Inflation time series were taken from Haver Analytics.

<sup>17</sup>Since respondents to *Consensus Forecasts* vary from country to country, these euro area constructs are not, strictly speaking, forecasts for the euro area economy. However, unless respondents of a particular country have systematic biases in their inflation forecast, the average inflation forecast across countries should be close to a forecast by an “average” forecaster for the average of the countries, that is, for the euro area as a whole.

<sup>18</sup>These two statistics are the most frequently used statistics to evaluate our sample forecasting performance.

## REFERENCES

- Banerjee, A., and M. Marcellino**, 2003, "Are there any reliable leading indicators for U.S. inflation and GDP growth?," Innocenzo Gasparini Institute for Economic Research, working paper, No. 236, April.
- Banerjee A., M. Marcellino, and I. Masten**, 2003, "Leading indicators of euro area inflation and GDP growth," Center for Economic Policy Research, discussion paper, No. 3893, May.
- Brave, Scott, and Jonas D. M. Fisher**, 2004, "In search of a robust inflation forecast," *Economic Perspectives*, Federal Reserve Bank of Chicago, Vol. 28, No. 4, Fourth Quarter, pp. 12–31.
- Castelnuovo, E., S. Nicoletti-Altimari, and D. Rodríguez-Palenzuela**, 2003, "Definition of price stability, range, and point inflation targets: The anchoring of long-term inflation expectations," in *Background Studies for the ECB's Evaluation of its Monetary Policy Strategy*, O. Issing (ed.), Frankfurt, Germany: European Central Bank, pp. 43–90.
- Dhyne, E., L. Álvarez, H. Le Bihan, G. Veronese, D. Dias, J. Hoffmann, N. Jonker, P. Lünemann, F. Rumler, and J. Vilmunen**, 2005, "Price setting in the euro area: Some stylized facts from individual consumer price data," European Central Bank, Eurosystem Inflation Persistence Network, working paper, No. 524, September.
- Estrella, A., and J. Fuhrer**, 1999, "Are 'deep' parameters stable? The Lucas critique as an empirical hypothesis," Federal Reserve Bank of Boston, working paper, No. 99-4, September.
- Fabiani, S., M. Druant, I. Hernando, C. Kwapil, B. Landau, C. Loupiaz, F. Martins, T. Mathä, R. Sabbatini, H. Stahl, and A. Stockman**, 2005, "The pricing behavior of firms in the euro area: New survey evidence," European Central Bank, Eurosystem Inflation Persistence Network, working paper, No. 535, October.
- Fisher, Jonas D. M., Chin Te Liu, and Ruilin Zhou**, 2002, "When can we forecast inflation?," *Economic Perspectives*, Federal Reserve Bank of Chicago, Vol. 26, No. 1, First Quarter, pp. 30–42.
- Goodfriend, M.**, 2007, "How the world achieved consensus on monetary policy," *Journal of Economic Perspectives*, Vol. 21, No. 4, Fall, pp. 47–68.
- Gürkaynak, R., A. Levin, and E. Swanson**, 2006, "Does inflation targeting anchor long-run inflation expectations? Evidence from long-term bond yields in the U.S., UK, and Sweden," Bilkent University, Ankara, Turkey, working paper, March 1, available at [www.bilkent.edu.tr/~refet/Gurkaynak\\_Levin\\_Swanson\\_2006mar01.pdf](http://www.bilkent.edu.tr/~refet/Gurkaynak_Levin_Swanson_2006mar01.pdf).
- Issing, O. (ed.)**, 2003, *Background Studies for the ECB's Evaluation of its Monetary Policy Strategy*, Frankfurt, Germany: European Central Bank.
- Labhard, V., G. Kapetanios, and S. Price**, 2007, "Forecast combination and the Bank of England's suite of statistical forecasting models," Bank of England, working paper, No. 323, May.
- Levin, A. T., F. M. Natalucci, and J. M. Piger**, 2004, "Explicit inflation objectives and macroeconomic outcomes," European Central Bank, Eurosystem Inflation Persistence Network, working paper, No. 383, August.
- Lindé, J.**, 2001, "The empirical relevance of simple forward- and backward-looking models: A view from a dynamic general equilibrium model," Sveriges Riksbank, working paper, No. 130, December.
- Maćkowiak, B., and M. Wiederholt**, 2005, "Optimal sticky prices under rational inattention," Humboldt University, discussion paper, No. 2005-040, August 4, available at <http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2005-040.pdf>.
- Mankiw, G., and R. Reis**, 2002, "Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve," *Quarterly Journal of Economics*, Vol. 117, No. 4, November, pp. 1295–1328.
- Roger, S., and M. Stone**, 2005, "On target? The international experience with achieving inflation targets," International Monetary Fund, working paper, No. WP/05/163, August.

**Rudd, J., and K. Whelan**, 2006, “Can rational expectations sticky-price models explain inflation dynamics?,” *American Economic Review*, Vol. 96, No. 1, March, pp. 303–320.

**Sargent, T.**, 1993, *Bounded Rationality in Macroeconomics*, Oxford: Clarendon Press.

**Sims, C.**, 2003, “Implications of rational inattention,” *Journal of Monetary Economics*, Vol. 50, No. 3, April, pp. 665–690.

**Stock, J., and M. Watson**, 2003, “Forecasting output and inflation: The role of asset prices,” *Journal of Economic Literature*, Vol. 41, No. 3, September, pp. 788–829.

**Svensson, L.**, 1999, “Inflation targeting as a monetary policy rule,” *Journal of Monetary Economics*, Vol. 43, No. 3, June, pp. 607–654.

**Vermeulen, P., D. Dias, M. Dossche, E. Gautier, I. Hernando, R. Sabbatini, P. Sevestre, and H. Stahl**, 2007, “Price setting in the euro area: Some stylized facts from individual producer price data,” European Central Bank, Eurosystem Inflation Persistence Network, working paper, No. 727, February.