

IT'LL ONLY HURT A SECOND?
MICROECONOMIC DETERMINANTS
OF WHO GETS FLU SHOTS

John Mullahy

Working Paper **6500**

NBER WORKING PAPER SERIES

IT'LL ONLY HURT A SECOND?
MICROECONOMIC DETERMINANTS
OF WHO GETS FLU SHOTS

John Mullahy

Working Paper 6500
<http://www.nber.org/papers/w6500>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 1998

I am grateful to Lee Schmeltz for supplying supplementary data used in the analysis and to Anita Chawla, Susan Ettner, Jon Gruber, Will Manning, Tom Philipson, Doug Staiger, and participants in seminars at Wisconsin, the BU/Harvard/MIT Health Economics Workshop, and the 1997 AEA Meetings in New Orleans for helpful comments and discussions. This research has been supported in part by a grant from the David and Lucile Packard Foundation. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

© 1998 by John Mullahy. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

It'll Only Hurt a Second? Microeconomic
Determinants of Who Gets Flu Shots
John Mullahy
NBER Working Paper No. 6500
April 1998
JEL No. I1

ABSTRACT

Appreciating how propensities to be immunized against the flu depend on individual characteristics and environments is essential if policies regarding influenza control are to be sensibly formulated. Beyond epidemiology, there are some important economic issues that must be addressed if the determinants of this form of preventive care are to be comprehensively understood. One concerns the relationship between labor supply and the propensity to be immunized: While it is costly (in terms of time costs) for workers to obtain immunizations, it is also workers who are likely to have relatively most to lose from being ill with the flu. Another concern not generally appreciated is the extent to which individuals' perceived risks of infection may affect their propensities to be immunized.

The analysis is based on data from the 1991 National Health Interview Survey. Immunization propensity displays the expected patterns by age and health status, while the results with respect to race, household structure, income and insurance are somewhat more surprising and/or novel. The estimated labor supply and perceived risk effects suggest that some aspects of the economics of preventive care generally not considered in empirical work are -- at least in this application -- important and merit further consideration.

John Mullahy
Departments of Preventative Medicine and Economics
University of Wisconsin-Madison
Madison, WI 53706
and NBER
jmullahy@facstaff.wisc.edu

I. INTRODUCTION

Each year in the U.S. anywhere from 10 to 50 million individuals become ill with the flu. In a typical year, about 20,000 of these individuals -- 90% of whom are elderly -- die from complications of the disease (Zamula, 1994). During influenza seasons, significant excess hospitalizations among the elderly have been identified (McBean et al., 1993). Upwards of 10% of physician office visits during peak flu season may be attributable to influenza-like illnesses (Chapman et al., 1993). The value of time lost at work, at school, and in household production due to influenza-related illness is considerable (U.S. OTA, 1981). While the costs associated with these and related phenomena are clearly substantial (U.S. OTA, 1981; McBean et al., 1993), it is also probably obvious that they fail to account for the additional real costs associated with the misery suffered each year by millions of flu victims.

Since for at least some large sub-populations influenza immunization has been demonstrated to be cost-effective (U.S. OTA, 1981), it is useful to understand more clearly why it is that only about one-seventh of the adult population receives a flu shot in any particular year. However, apart from the observed regularity that older individuals and those in other high-risk groups (the infirm, health care workers, etc.) are more inclined than others to be immunized, the literature offers little evidence about the other main determinants of influenza immunization. Such knowledge would be particularly useful since policies and strategies to increase influenza immunization rates are actively under consideration (ACIP, 1993).

The main objective of this paper is to analyze empirically the economic determinants of why some individuals obtain flu shots

while others don't. While the empirical results are entirely consistent with what is commonly known or believed about influenza immunization -- i.e. that the elderly and the infirm have relatively high propensities to be immunized -- the main contribution of the paper is in demonstrating the importance of other possibly less obvious factors.

Some fundamental economic issues must be addressed if the determinants of this form of preventive care are to be comprehensively understood. One of these concerns the relationship between labor supply and the propensity to be immunized. There has been some considerable recent debate about the cost-effectiveness of immunization of healthy, non-elderly adults (Nichol et al., 1995). Understanding how labor market behavior affects immunization propensities is, therefore, an important consideration in such cost-effectiveness calculations. While it is relatively costly (in terms of time costs) for workers to obtain immunizations, it is also these workers who are likely to have relatively most to lose from being ill with the flu. Using instrumental variables methods, the paper attempts to circumvent some possibly spurious correlations to address this issue.

Another concern not generally appreciated is the extent to which individuals' perceived risks of infection may affect their propensities to be immunized. The effects of such risk perceptions on defensive behavior have been identified empirically for in such diverse contexts as HIV infection (Ahituv et al., 1996) and measles infection (Philipson, 1996). Using information on the severity of the flu season prior to the season during which the individual made his/her decision to obtain a flu shot, the paper attempts to shed some light on this issue.

The plan of the paper is as follows. Section II presents some background information on influenza immunization in the U.S. that is germane to the empirical analysis conducted in section V.

Section III describes a simple economic framework of the demand for preventive care that generates some predictions about which individuals will be most inclined to be immunized. Section IV discusses the data used in the empirical analysis and discusses the empirical strategy. The data are from the 1991 National Health Interview Survey (NHIS), supplemented by information on flu season characteristics and labor market conditions in individuals' states of residence. Sections V and VI present the empirical findings. Section VII summarizes.

II. SOME BACKGROUND

To provide context for the analysis it is useful to provide some background information. While influenza immunization is a relatively uncomplicated form of preventive care, there are some natural and institutional aspects to flu shot administration that are important to appreciate if the empirical analysis is to be sensible and useful.

There is mounting empirical evidence on the effectiveness and cost-effectiveness of influenza vaccination (U.S. OTA, 1981; Govaert et al., 1994). While two promising antiviral drugs, amantadine and rimantadine, have been developed and are in some use, it is generally thought that immunization -- though not 100% effective -- is more effective and more cost-effective than antivirals for preventing the adverse health consequences of flu infections. Special consideration in targeting programs to enhance immunization rates has been given to individuals in those groups who are at greatest risk for suffering serious flu-related

complications. Most prominent in these target groups are the elderly (ages 65+) and non-elderly individuals in poor health and/or immunocompromised.¹

The composition of each year's trivalent² vaccine evolves annually with the "best guess" recommendations made early each year by the FDA's Vaccines and Related Biologicals Advisory Committee. With a good match between the vaccine's antigens and those actually in circulation, the Advisory Committee on Immunization Practices (ACIP) estimates that about 70% of illness in non-elderly populations and of hospitalizations of elderly individuals residing outside chronic-care facilities is prevented (ACIP, 1994).

While "flu season" runs officially from October to April in the U.S., influenza outbreaks are most prominent from late December to early March (ACIP, 1994). Each year's flu vaccine becomes available for administration in September, with the recommended optimal time for immunization of targeted high-risk groups (the elderly, the infirm, health care workers, etc.) being mid-October to mid-November. Immunizing too early can result in

1. ACIP, 1994. Additionally:

Influenza vaccine is strongly recommended for any person ≥ 6 months who -- because of age or underlying medical condition -- is at increased risk for complications of influenza. Health-care workers and others (including household members) in close contact with persons in high-risk groups should also be vaccinated. In addition, influenza vaccine may be administered to any person who wishes to reduce the chance of becoming infected with influenza. (ACIP, 1994)

2. Each year's vaccine typically contains two strains of Influenza Type A virus and one strain of Influenza Type B virus.

diminished antibody levels by the time peak flu season arrives, while immunizing too late may not allow for the two-to-four weeks required for building antibody concentrations to a level sufficient to combat exposure to the virus (Zamula, 1994). Most flu shots in the U.S. are administered between October and December, although official government recommendations suggest that the vaccine be offered to patients in outpatient settings beginning in September and continuing throughout the peak flu season (ACIP, 1994).

As will be discussed in greater detail below in section IV, the data analyzed in this paper are from a survey conducted during 1991, with the key information elicited being whether the respondent reports having received a flu shot in the twelve months prior to the survey. As such, it should be recognized that for the majority of the survey respondents this response will refer to immunizations received (or not received) during the 1990-91 flu season, with the most prominent activity being October-December 1990. Given the scheduling of influenza immunizations, only those respondents questioned late in 1991 would have any likelihood of having received their shots in 1991 (the data do not contain information on precisely when the shot was obtained); because of this possible confounding, observations on these individuals will be dropped from the empirical analysis.

One issue addressed below is whether individuals' propensities to obtain flu shots may be affected by their experience during the most recent flu season. As such, for most survey respondents analyzed, the most recent flu season will be the 1989-90 flu season, although for those surveyed late in 1991 the 1990-91 flu season may be the relevant flu season.

It turns out that the 1989-90 and the 1990-91 flu seasons were quite different. The CDC reported the 1989-90 flu season to be "slightly shorter and more intense" than the 1990-91 season (Chapman et al., 1992). During the 1989-90 flu season, 30 of the 50 states plus D.C. reported at least one week of "widespread" flu activity at some point during the flu season, while only 17 states reported such widespread activity during the 1990-91 season.³

Another consideration is that from 1988 to 1992 Medicare conducted demonstration projects in parts of nine states (AZ, IL, MA, MI, NY, NC, OH, PA, TX) and in the entire state of Oklahoma to assess whether and, if so, by how much Medicare insurance coverage of flu shots and other enhancements in the vaccine delivery system might increase vaccine uptake and improve health outcomes in the Medicare population (Schmitz et al., 1993). The results indicated that the demonstrations "increased annual influenza vaccine coverage and measured both health and economic benefits of influenza vaccine for Medicare" (MMWR 8/13/93 (Editorial)). The main upshot of the demonstration projects was that flu shots became a covered Medicare benefit in May 1993, with full coverage when the vaccine is administered by a Medicare-participating provider (Marwick, 1994). While the survey data available here do not indicate whether any particular elderly individual was involved in one of the demonstration projects, it may be that

3. These data are from the CDC's Weekly State and Territorial Reports on flu activity, as assessed and reported by the state and territorial epidemiologists. "Widespread" flu activity is said to occur when there are "outbreaks of influenza-like illness or culture-confirmed influenza in counties having a combined population of at least 50% of the state's population." These unpublished data were graciously provided to me by Lee Schmeltz of CDC's National Center for Infectious Diseases.

elderly residents of the demonstration project states on average had their immunization propensities affected in 1991 by the projects.

III. THE DEMAND FOR FLU SHOTS: A CONCEPTUAL FRAMEWORK

An individual's decision to be immunized when flu season is approaching is a more-or-less classic example of the demand for preventive care under uncertainty: the individual doesn't know whether or when he or she will be exposed to the infectious agent; doesn't know how long or with what severity the illness, if contracted, will last; and doesn't know whether preventive measures like immunization will be effective if undertaken. The decision, therefore, is a classic discrete choice in which the expected net welfare from being immunized is compared with the expected net welfare of not being immunized.

A stylized description of the decision process of concern here is presented in fig. 1. In the context of a simple two-period model, in which the decision of interest here is the period two choice of whether to obtain a flu shot, an individual can be thought to have made a period zero decision to obtain a flu shot.

Given this decision -- which, for reasons discussed above, takes place before the onset of flu season -- the individual's community either has or does not have an "outbreak" of flu (i.e. a severe flu season). Given the shot decision made by the individual and the severity of flu in the community, the individual will or will not get sick with flu. Finally, given past immunization behavior, past severity behavior, and prior illness experience, the individual makes the period one immunization decision. In this paper, it will be possibly to account empirically only for the prior year's "outbreak" severity and the current year's

immunization decision; the other components of this decision process are not observed in the data.

A straightforward way to characterize the economic aspects of this decision problem is to begin by specifying a health production function

$$H_t = H(S_t, S_{t-1}, H_{t-1}, E_t, M_t, \Omega_t),$$

indicating that the severity of influenza-related illness in year t (H_t) depends on whether the individual received a flu shot prior to the year t flu season (S_t), whether the individual received a flu shot or had the flu the previous year (S_{t-1}, H_{t-1}), the extent to which the individual is exposed to flu virus in year t (E_t), prescription and/or OTC medications used by the individual in year t (e.g. antiviral prescription medications, pain relief medications, etc.), and the levels of other exogenous covariates at time t (Ω_t) (e.g. age and other factors that might tend to exacerbate the severity of flu symptoms).

The data contain no information on whether the individual utilizes any form of medication, so this interesting consideration must henceforth be ignored. There is also no information on whether the individual obtained a flu shot in the previous year (S_{t-1}) or had the flu in the previous year (H_{t-1}). Apart from revealing potentially useful information about individuals' flu shot propensities, there is also some debate in the literature about whether immunization with or natural infection by a particular strain of flu virus in a prior or recent year might provide sufficient antibody response to mitigate

infection by an antigenically similar strain of flu in the present (Gill and Murphy, 1985; Hoskins et al., 1979). To the extent that such immunity does confer, then the expected "marginal product" of the present year's immunization would be reduced.⁴ With such data limitations recognized, the relevant version of the health production function for the empirical purposes at hand is

$$H_t = H(S_t, E_t, \Omega_t).$$

Individuals have essentially three choices during immunization season each fall: obtain a flu shot by taking time off from work (W) to obtain the shot ($S_W=1$, $=0$ otherwise); obtain a flu shot by giving up leisure time (ℓ) to obtain the shot ($S_\ell=1$, $=0$ otherwise); or don't obtain the flu shot ($S_N=1$, $=0$ otherwise).⁵

The individual will be assumed to make choice i if it yields the greatest expected utility, $EU(S_i|\mathbf{X}) > EU(S_j|\mathbf{X})$ and $EU(S_i|\mathbf{X}) > EU(S_k|\mathbf{X})$ for $i, j, k \in \{W, \ell, N\}$, where the uncertainty in the model is due to the probabilistic nature of flu-related health outcomes given any of the choices, and where \mathbf{X} summarizes all conditioning

4. This message is one not commonly conveyed in immunization campaigns where considerable emphasis is placed on the importance of repeated annual immunization. It is probably a reasonable conjecture that most individuals do not take into account such persistent immunity possibilities when making their annual immunization decisions, although such a conjecture is clearly refutable.

5. While the data do not indicate when and where a shot was obtained, it is nonetheless useful to describe two alternative possibilities for obtaining flu shots that many working individuals are likely to face.

covariates. The analysis can then be cast in a standard discrete choice econometric model by appending additive stochastic elements to the utility functions, giving $V(S_i|\mathbf{X})=EU(S_i|\mathbf{X})+\varepsilon_i$.

To simplify the analysis it is useful to dichotomize H_t into "gets the flu" (H_0) and "doesn't get the flu" ($H_1>H_0$) outcomes and to then analyze the problem in an expected utility context in which the individual considers how the probabilities of healthy (doesn't get the flu) and sick (gets the flu) outcomes vary depending on whether a flu shot is obtained and on other factors.⁶

That is, the health production function is a conditional probability model, with

$$\phi_j = \text{Prob}(H=H_1|S_j,\mathbf{X}), \quad j \in \{S,N\},$$

where $S_S=S_W+S_\ell$.

The health production/probability functions provide what might be thought of as the individual's prior probabilities on health outcomes. However, before making a decision on the basis of expected utility comparisons, a rational Bayesian decisionmaker

6. "Intermediate" health outcomes might be of interest as well. For example, after becoming ill with flu an individual might consider expending time and money to consult a physician and possibly obtain a prescription for an antiviral medication. Moreover, many individuals will purchase (or draw down existing stocks of) over-the-counter medications for relief of pain and other symptoms after they become ill. In these instances, it would be reasonable to suggest that the realized health outcomes are probably better than they would be in the absence of such medications but not as good as would be the case if flu symptoms were avoided altogether. Such issues -- while perhaps important in many real-world settings -- are ignored here.

will utilize observable data to update these prior probabilities. One form of readily-accessible data that might be used for such updates is the individual's recent experiences with the flu in the community, the workplace, or the household. It may be that the severity of the flu season in the prior year (\mathbf{F}) is decision-relevant information for the individual with, for instance, a higher level of severity making the individual more aware of the preventive benefits of immunization, resulting in an upward update of ϕ_S . Whatever the updating process, it is assumed that ϕ_S and ϕ_N are updated to π_S and π_N , respectively, with

$$\pi_j = \text{Prob}(H=H_1 | S_j, \mathbf{X}), \quad j \in \{S, N\},$$

and that π_S and π_N are the probabilities on which the individual bases his/her immunization decision (\mathbf{X} is henceforth assumed to contain \mathbf{F}).

It is assumed that data on \mathbf{F} update ϕ_S and ϕ_N to π_S and π_N in a manner such that $d(\pi_S - \pi_N)/d\mathbf{F} > 0$. Several reasonable possibilities would suggest such a result. A more severe prior flu season (higher \mathbf{F}) might make individuals more aware of the protective benefits of immunization without affecting their perceptions of the risks associated with being unimmunized ($d\pi_S/d\mathbf{F} > 0$, $d\pi_N/d\mathbf{F} = 0$). Conversely, individuals might respond to a more severe prior year outbreak by lowering their subjective probabilities of the healthy outcome both with and without immunization, but might revise π_N downward relatively more than π_S ($0 > d\pi_S/d\mathbf{F} > d\pi_N/d\mathbf{F}$). Other possibilities might be entertained as well. Data on the severity of flu outbreaks in the individual's

community (state of residence, to be more precise) are available and will be incorporated into the empirical analysis to follow.⁷

The main features of preference relevant for this discussion are captured in a simple model where individuals have instantaneous⁸ separable utility functions

$$U(H, \ell, S, Z) = \Psi(H) + \Lambda(\ell) - \beta S + Z,$$

where H is flu-related health status, ℓ is leisure time, $S=1-S_N$ is a 0-1 indicator of receiving a flu shot, and Z is composite consumption (time period subscripts are dropped henceforth unless required for clarification). It is assumed that $\Psi', \Lambda', \beta > 0$ and $\Psi'', \Lambda'' < 0$. β measures the disutility of receiving the shot itself:

7. Note that this setup differs in some respects from the "community incidence/prevalence" models considered by Ahituv et al., 1996, for HIV infection and by Philipson, 1996, for measles. For HIV, observations on recent patterns of the community's prevalence rate are suggestive of present-period infection risks for the individual. For measles, recent observations on the community's incidence/prevalence rate might suggest several possibilities (herd immunity; extraordinary risk; etc.) about present-period infection risks. For influenza, however, neither the herd immunity effect of measles infection nor the incurability of HIV infection would appear to be particularly relevant concerns.

8. A perhaps more suitable model of preferences would be a two-period model in which the decision to obtain a flu shot is made at $t=0$ but where flu-related health outcomes are realized at some later time $t=1$, where the gap between $t=0$ and $t=1$ might be on the order of three months. While in some sense a relatively short time interval, such a two-period formulation would nonetheless allow for considerations of discounting which could be important in this class of decisions to the extent that individuals have very high rates of time preference.

the instantaneous "ouch" factor, the generally mild and transient side-effects like soreness at the injection site, etc.⁹ Since influenza vaccine consists of inactivated virus (as opposed to a live virus vaccine like oral polio, for example), infection with influenza caused by immunization is generally not possible (contrary to much common opinion). Some central analytics of the model are presented in the Appendix.

IV. DATA AND EMPIRICAL STRATEGY

1991 NHIS Sample

Most of the data are from the Health Promotion and Disease Prevention (HPDP) Supplement to the 1991 National Health Interview Survey. The NHIS survey is administered annually to individuals in approximately 50,000 households, or roughly 140,000 individuals (U.S. National Center for Health Statistics, 1989). The structure of the survey is a probability sample of the noninstitutionalized population, and the recent versions of the NHIS have been structured to oversample blacks in order to obtain more precise estimates of health-related characteristics of the black population. Each year the NHIS obtains information on household demographics and household members' health status. The survey is conducted during the entire calendar year, so that individuals in all seasons are surveyed; some implications of this for the present analysis are discussed below.

9. The common side effects that have been identified include soreness, fever, myalgia, malaise, and possible allergic reactions to the egg protein used to formulate the vaccine (American College of Physicians, 1990).

The variables used in the analysis are described and summarized in tables 1 through 3. The dependent variable, FLU SHOT, is created from the response to the survey item:

*"During the past 12 months, have you had a flu shot?"
(Read if necessary: "This vaccination is usually given
in the fall and protects against influenza for the flu
season.")*

Rationales for the inclusion of most of the covariates are obvious, but in some cases further discussion is appropriate at this juncture.

The health status and age (linear and quadratic) variables are included to describe individuals in primary target groups for immunization. The existing literature would suggest that these characteristics are likely to be the dominant determinants of flu shot propensities. The quadratic age profile is specified to allow for a flexible shape for the age-related takeup rate of immunizations.¹⁰ The self-perceived health status variables (VERY GOOD,...,POOR, with excellent health status omitted) are included to test the hypothesis that the better an individual's self-perceived health status, the lower will be that individual's propensity to be immunized, all else being equal.

Second, each individual's health insurance coverage status is determined from the 1991 NHIS Family Resources supplement. The insurance measure constructed for the analysis (INSURED) is

10. Given the age-based subsamples used here (25-64 and 65+), this appears to provide a satisfactory fit to the data. As suggested below, however, a *quartic* age profile may be a more suitable characterization of the age-immunization profile for the entire pooled sample of ages 25+.

admittedly rough. It is not expected that the parameter estimate associated with this covariate will be picking up a pure effect of flu shot coverage on the propensity to receive a flu shot.¹¹ Instead, INSURED is more appropriately interpreted as a very broad measure of access to the health care system, only one part of which might represent actual coverage for influenza vaccine and its administration. Additionally, the interpretation of this insurance effect is one that holds constant (in some specifications) labor supply behavior.¹²

Two measures of labor supply are used in the empirical work. WORKS is a binary indicator of the individual's usual activity. HOURS is a measure of the number of hours an individual typically works each week. Additionally, a measure of log-wage (LWAGE) is used in an attempt to assess how the value of time might affect immunization propensities.

Because of their above-average rates of contact with infected individuals, health care workers are a designated high-risk group for whom immunization is recommended. To control for this risk factor without introducing intractable endogeneity problems, a variable HEALTHCARE WORKER is created that indicates whether any adult in the household works in a health care occupation.¹³

11. In 1991, Medicare coverage had not yet been implemented (except in the Demonstration Project). Many private insurance plans do cover flu shots for individuals who are determined to be in the "high risk" categories described earlier.

12. The potential endogeneity of insurance status is not addressed here. Since the insurance status measure is quite rough, self-selection into insurance would not appear to be a major issue here.

13. This coding corresponds to NHIS occupational recode 1 equal to 07 or 08.

Finally, the WIDESPREAD FLU variable, measuring the number of weeks during the 1989-90 flu season the individual's state of residence was deemed to have widespread influenza outbreaks, is included for the "perceived risk/effectiveness" reasons discussed in section III,¹⁴ and is analogous to the empirical approach used by Philipson, 1996, in studying measles immunization propensities.¹⁵ This measure, based on information obtained from CDC's Weekly State and Territorial Reports on flu activity for the 1989-90 flu season¹⁶ is merged to the NHIS file by the individual's state of residence.¹⁷

14. The conceptual model described in section III suggested that individuals' propensities to receive flu shots might depend on their perceived risks of being infected with flu viruses. Using a proxy like WIDESPREAD FLU for the previous season's severity would tend to capture the effects of such perceived risks. However, it should be noted that there might also be some "supply side" effects due to a prior season's severity, e.g. health care providers and local health departments might respond to a particularly severe flu season by encouraging patients/residents to receive flu shots during the subsequent immunization season. As such, the data available here will not be able to identify whether any significant estimated relationship is a demand-side effect, a supply-side effect, or some combination thereof.

15. Philipson, 1996, finds a strong positive association between lagged measles incidence and propensities to be immunized against measles.

16. The distribution of the WIDESPREAD FLU severity measure across the states is as follows:

WIDESPREAD FLU (# Weeks)	States

0	AL, AR, CA, CO, FL, HI, IL, IN, KS, ME, MD, NV, NJ, NM, OR, PA, RI, SC, VT, WV
1	DE, DC, KY, NH, TN
2	LA, OK, WY

(continued)

Sample Definitions

The empirical analysis proceeds for the most part by splitting the pooled sample by age, considering a non-elderly subsample (ages 25-64) and an elderly subsample (ages 65+). Such an age-based subsampling scheme would be suggested by the ACIP recommendations targeting elderly individuals for immunization. While the HPDP surveys individuals ages 18+, the minimum age is set here at 25 years in order to obtain a sample of respondents who will (for the most part) have completed their schooling.¹⁸ For the reasons discussed above, to enhance the likelihood that

3	CT, MI, MN, OK, SD, WA
4	MA, MS, MO, UT
5	AK, GA, MT, NC, VA
6	ID, IA, WI
7	AZ, TX
8	--
9	NY

17. Of the 43,732 observations in the HPDP Supplement, 42,971 (98.3%) were on individuals responding either "yes" or "no" to the flu shot question. The missing data comprise 578 observations (1.3%) for which the flu shot information was "not ascertained," and 183 observations (0.4%) with either "don't know" or "refused" responses. A simple logit regression of the probability of unavailable flu shot data revealed that data unavailability is positively correlated with being male, being nonwhite, and having relatively low schooling attainment; controlling for these other factors, there was no statistically significant association between flu shot data availability and age. The analysis will, henceforth, ignore the observations with the missing flu shot data; accordingly, inferences should be made cognizant of this sample selection.

18. This is to circumvent the issue of whether attained schooling or *ultimately attained* schooling is the more important determinant.

the analysis is focusing on flu shot behavior during a single flu season, the sample will additionally be restricted to individuals responding to the survey in the first three quarters of 1991. These exclusion criteria and the exclusion of a small number of observations due to missing data result in an estimation sample of 27,135 observations in the pooled sample. The age and sex subsample data on flu shots are presented in table 3.

Empirical Strategy

Let y be the 0-1 indicator corresponding to FLU SHOT. Dropping observation subscripts to minimize notational clutter, suppose the probability that $y=1$ conditional on covariates Ξ is given by the distribution function $G(\Xi\alpha)$, where α are the unknown parameters of interest. Let the observed k -vector of covariates X have corresponding parameters β and let the net effect of unobserved covariates on $\Xi\alpha$ be denoted Θ , i.e. $G(\Xi\alpha)=G(X\beta+\Theta)$. Then

$$\text{Prob}(y=1|X,\Theta) = G(X\beta + \Theta),$$

so it follows that

$$y = G(X\beta + \Theta) + \varepsilon$$

where

$$E(\varepsilon|X,\Theta) = 0$$

by construction, implying as well that $E(\varepsilon|X)=0$ by iterated expectations.

The main complication of interest here is the possibility that X and Θ are correlated in the particular sense that $E(\Theta|X)=h(X)$ is a nontrivial function of X . That is, some of the unobserved influences on the probability of obtaining a flu shot may be systematically correlated with some of the observed covariates.

Of particular concern is the likelihood that individuals' labor market behaviors will be governed in part by unobserved factors that may also be important determinants of their propensities to obtain flu shots, e.g. unobserved attributes of health/disability status, time preference, risk aversion, etc. That is, giving $G(\cdot)$ the interpretation of a structural econometric demand function for flu shots, its conceptual counterpart might be thought of as a conditional (on labor supply) demand function (Browning and Meghir, 1991). As such, standard parametric estimation strategies like probit and logit will not generally provide consistent estimates of the parameters of interest, β , so long as the unobservables do jointly determine immunization propensities and labor supply behavior.

Since only one cross-section sample is available, the only apparent solution to this problem in the present context is an instrumental variable (IV) approach. As argued below, conceptually reasonable instruments, Z , are available (Z is a $(p \times 1)$ vector with $p \geq k$). Specification of $G(\cdot)$ thus becomes the central issue. Probit and logit are possibilities, but IV estimation in these cases requires specification and consistent estimation of the reduced forms of the suspected "endogenous" variables as well as adjustment of estimated covariance matrixes for inclusion of predicted regressors. This entire process is also likely to be non-robust to misspecification of the

probability model $G(\cdot)$ (e.g. logit is truth but probit is estimated), the functional form of the reduced forms $X=X(Z)$, or both.¹⁹

The estimation strategy adopted here is suggested by the work of Heckman and MaCurdy, 1985, and White, 1982. The basic idea underlying this estimator is to specify what amounts to a linear probability model for $y|\Xi$, i.e. $G(\cdot)=(\cdot)$, and then to use an IV estimator to circumvent the correlation between X and Θ . That is,

$$\begin{aligned} y &= \Xi\alpha + \varepsilon \\ &= X\beta + \Theta + \varepsilon \\ &= X\beta + u, \end{aligned}$$

where $E(u|X) \neq 0$ since $E(\Theta|X) \neq 0$. It is assumed that $E(u|Z)=0$ (i.e. instruments uncorrelated with unobservables) and that $\lim(Z'X/N)$ is of rank k (i.e. instruments correlated with observed covariates). In this case the IV estimator is closed-form and is given by²⁰

$$\hat{\beta} = [X'Z(Z'\hat{\Sigma}Z)^{-1}Z'X]^{-1}X'Z(Z'\hat{\Sigma}Z)^{-1}Z'y,$$

where the preliminary estimate $\hat{\Sigma}$ is obtained here as $\hat{\Sigma}=\hat{\varepsilon}\hat{\varepsilon}'$ using 2SLS residuals $\hat{\varepsilon}$. The asymptotic covariance matrix can be estimated consistently by $\hat{V}(\hat{\beta})=[X'Z(Z'\hat{\Sigma}Z)^{-1}Z'X]^{-1}$

19. For comparison purposes, however, results for such a probit specifications will be presented below in a footnote.

20. The N -row vectors and matrixes corresponding to y , X , Z , and ε are denoted in boldface.

In the overidentified case, Hansen, 1982, has shown that the GMM approach provides a straightforward test of overidentifying restrictions (known often as the Hansen's J-test), as the value of the criterion function evaluated at $\hat{\beta}$ has asymptotically a χ^2 distribution with p-k degrees of freedom under the null hypothesis of orthogonality between ε and Z.

V. EMPIRICAL RESULTS

The variable definitions are summarized in table 1 and sample means are presented in table 2. Table 3 displays a breakdown of sample FLU SHOT means by age and gender. Within age categories, there is little apparent gender variation in these marginal means. Conversely, but not unexpectedly, large age differences in immunization behavior are obvious for males and females.

Baseline Results

The baseline OLS linear probability results are presented in table 4. Columns 1-3 and 4-6 present the results for the non-elderly sample (ages 25-64) and the elderly sample (ages 65+), respectively. The specifications in columns 1 and 4 contain no labor supply variables; those in columns 2 and 5 include the WORKS measure; and those in columns 3 and 6 include the HOURS measure. For later purposes, it should be noted here that all these specifications include in the covariate vectors the FAMILY SIZE and MARRIED variables.

While the main focus after table 3 will be on the non-elderly sample, the estimates in table 3 are in and of themselves useful in depicting the basic covariance structures underlying immunization propensities for both elderly and non-elderly. Some

of the findings will be discussed further in section VII insofar as their policy implications are concerned.

For purposes of discussion, it is useful to group the covariates into three categories: labor market variables (WORKS, HOURS, LWAGE); health status (VERY GOOD, GOOD, FAIR, POOR); other risk factors (WIDESPREAD FLU, HEALTH CARE WORKER, AGE); and other demographics (MALE, SCHOOLING, NONWHITE, INSURED, FAMILY SIZE, MARRIED, and the census region and MSA dummies).

Widespread Flu

The lagged WIDESPREAD FLU point estimates are positive and are statistically significant for both the non-elderly and the elderly, with the magnitudes of the estimated effects considerably larger for the latter. Thus, as hypothesized, individuals' propensities to obtain flu shots do appear to respond to recent experiences with outbreaks in their communities, a finding consistent with those obtained by Ahituv et al., 1996, for HIV infection and Philipson, 1996, for measles. As indicated in section III, such a finding might arise because of several (noncompeting) explanations.

Since California was one of the states reporting zero weeks of WIDESPREAD FLU severity and since New York was the state reporting the greatest duration of severe flu (nine weeks), there was some concern that the estimates might be unduly influenced by various unobserved features of these two highly-populated states.

Accordingly, the baseline models in columns 1 and 4 of Table 3 were reestimated on samples that excluded observations from California and New York (N=17,081 for ages 25-64; N=5,024 for ages 65+). The resulting OLS estimates (t-statistics) are .00242 (2.7) versus .00164 (2.4) in the full 25-64 sample, and .00966 (3.4)

versus .00844 (3.8) in the full 65+ sample. There is clearly little evidence, then, that observations from these states serve in any meaningful way as influential outliers.

Another inferential issue arises when state-level data (like WIDESPREAD FLU) are merged with microdata records. If there are area-specific unobservables (e.g. state, PSU, or some other cluster-level fixed effects), then inferences that fail to account for the cross-observation correlation of unobservables within such clusters may be misleading (Borjas and Sueyoshi, 1994). To assess the implications of this possibility in these data, the baseline specification in column 1 of Table 3 was reestimated, this time using the Stata 3.1 procedure `hreg` with the `group(.)` option invoked. This gives a White-type heteroskedasticity-consistent covariance matrix but with a presumed block-diagonal Σ in $(\mathbf{X}'\mathbf{X}(\mathbf{X}'\Sigma\mathbf{X})^{-1}\mathbf{X}'\mathbf{X})^{-1}$, where the block boundaries correspond to the group definitions. Two covariance estimates were computed: when the group was specified to be the individual's PSU of residence, the corresponding t-statistic on WIDESPREAD FLU was 1.7; when the group was specified to be the individual's state of residence, the corresponding t-statistic was 1.1. As such, area-specific unobservables probably exert some influence on flu shot propensities.

Health Status

Self-perceived health status has essentially the same pattern of effects on both subsamples. As health status worsens from excellent (the omitted baseline) to poor, the propensity to be immunized increases sharply: Individuals in poor health are about 10% (non-elderly) or 17% (elderly) more likely than otherwise comparable individuals in excellent health to obtain a flu shot.

In part such behavior is consistent with a rational response to the fact that the consequences of influenza-related illness are likely to be far more severe in already-unhealthy individuals than in their healthy counterparts. It is also true that individuals in poor health are relatively more likely to have contacts with health care providers during the year, such contacts affording opportunities to obtain immunizations that might not otherwise arise. Some additional policy implications of these findings are discussed below.

A rather obvious point should be stressed: a simple self-reported general health status measure is unlikely to pick up all the key dimensions of how health status affects immunization propensities. In order to obtain a more comprehensive picture of how health status affects immunization propensities, some information on chronic respiratory conditions available from the 1991 NHIS condition file was merged. Individuals having chronic respiratory conditions are in one of the designated high risk groups for which immunization is most strongly encouraged. The NHIS administers randomly to each respondent one of six possible detailed probe surveys on chronic conditions. For the one-sixth of the sample to whom the respiratory condition survey was administered, an individual was classified as having a chronic respiratory condition (RESP=1) if they self-reported one or more of the Recode #3 chronic respiratory conditions (recodes 601-615); otherwise, RESP was set equal to zero for the other adults receiving the respiratory condition survey.²¹ The specifications

21. The prevalence rates turn out to be quite large (.36 in the non-elderly sample (N=3,467) and .31 in the elderly sample (N=1,007)), this attributable in part to the fact that chronic
(continued)

reported in columns 1-4 in table 4 were reestimated on these smaller samples and with the RESP variable included in the covariate vector. While the relatively small sample sizes result in substantially larger standard errors, most of the key inferences associated with the other parameters reported in table 4 remain intact.²²

In the specification comparable to that reported in column 1 of table 4, the point estimate of the RESP parameter is .035 ($t=3.4$), which can be compared -- to obtain a sense of relative magnitudes -- with the corresponding point estimates of the FAIR and POOR parameters of .044 ($t=1.9$) and .078 ($t=2.0$), respectively. An interpretation is that after controlling for general self-perceived health status, non-elderly individuals with chronic respiratory conditions are still about four percentage points more likely to obtain immunizations than others.²³

sinusitis is one of the most common chronic illnesses among adults in the U.S.

22. For the non-elderly sample, notable differences are that the WIDESPREAD FLU point estimate becomes $-.00014$ ($t=.08$) in the specification comparable to that reported in column 1, and that the point estimates on WORKS and HOURS (comparable to columns 2 and 3, respectively) are now positive but insignificant: $.0056$ ($t=0.5$) and $.00007$ ($t=0.3$), respectively. The positive point estimates owe to the sample composition, not to the inclusion of the RESP covariate, since the WORKS and HOURS point estimates and associated t -statistics are roughly the same in this sample whether or not RESP is included.

23. When both RESP and an interaction term $RESP*POOR$ are included, the results are even more striking: the main effect of POOR becomes insignificant ($-.026$ ($t=0.7$)), the main RESP effect is still significant ($.029$ ($t=2.8$)), and the interaction effect is extremely large and significant ($.197$ ($t=2.8$)). There are 100 observations in this sample for which $POOR=1$, 46 coinciding with $RESP=0$ and 54 coinciding with $RESP=1$.

For the elderly sample, the effects of chronic respiratory conditions on immunization propensity are even more pronounced. In the specification comparable to that reported in column 4 of table 4, the point estimate of the RESP parameter is .089 ($t=2.6$), which can be compared with the corresponding point estimates for FAIR and POOR of .035 ($t=0.7$) and .096 ($t=1.3$), respectively.²⁴

Other Risk Factors

The presence of a health care worker in the household is associated with an increased propensity to obtain immunization for both subsamples, with the estimated effect statistically significant in the non-elderly sample. Such results may arise due to a rational response to the higher likelihood of exposure to infection in such households, and is also consistent with the recommended immunization guidelines for at-risk groups. It may also be consistent with the possibility that health care workers may have relatively low time costs of obtaining flu shots (e.g. mass immunizations performed on site each fall in hospitals, medical schools, etc.).²⁵

The age pattern of immunization propensities is, as expected, monotonically increasing, convex for the non-elderly sample and

24. The RESP*POOR interaction was statistically insignificant in the specification that included this covariate for the elderly sample.

25. When an indicator of whether the individual him/herself was a health care worker was included instead of the household-level measure, the estimated magnitudes and significance levels were even greater. However, because of the greater possible endogeneity problems associated with such an inclusion, the household-level measure will be used in the remaining specifications to be considered.

concave for the elderly sample. The overall higher take-up rates for the elderly group are, of course, consistent with published recommendations for at-risk populations. Indeed, the rate of take-up is greatest around age 65, just as might be expected from the published recommendations: a linear probability regression of FLUSHOT on AGE, AGE², AGE³, and AGE⁴ using the pooled sample of individuals ages 25+ (N=27,135) revealed an inflection point in the age profile at between 65 and 66 years of age (see fig. 2). As such, defining the subsamples of ages 25-64 and ages 65+ effectively splits the pooled sample into a convex age-immunization profile subsample and a concave age-immunization profile subsample.

Other Demographics

The estimated effects of SCHOOLING on immunization propensities are positive and statistically significant, particularly so for the elderly sample where each additional year of schooling is associated with over a 1.5 percentage point *cet. par.* increase in immunization propensities. While statistically important and substantively meaningful, the literature clearly suggests that such large education effects are to be expected when examining preventive care behavior. Whether these estimated education effects are causal or simply a marker for unobserved factors like time preference is beyond the scope of this paper to consider.

The estimated effects of race are interesting. In the non-elderly sample, nonwhites are marginally more likely to obtain immunizations, all else equal, although the point estimates are not statistically significant. However, in the elderly sample the direction of the estimated effect is reversed and it is highly

significant.²⁶ The estimated insurance effects are positive and significant for the non-elderly sample, as might be expected when taking a "health care access" perspective. However, it should be recognized that virtually the entire elderly sample reports insurance coverage (i.e. all but 19 observations), so that not too much should be put on the point estimates in this case.

FAMILY SIZE and MARRIED have small and statistically insignificant effects on immunization propensities in the non-elderly sample, a result that comes into play in an important way when instrumental variable strategies are considered below. In the elderly sample, however, being married is associated with a very large (7.5 percentage points) and statistically significant *cet. par.* increase in immunization propensities, a not-surprising finding that might be ascribed to any number of noncompeting explanations.

The estimated residential location parameters indicate that individuals residing in the south and west census regions are *cet. par.* relatively more likely than those residing in the northeast and midwest to obtain immunizations, and that individuals residing outside MSAs are *cet. par.* relatively more likely than those residing in MSAs to obtain immunizations.

Since the effects of these demographics are of secondary interest, and since they tend (with the exception of MALE) to be relatively insensitive to changes in specification or estimation strategy, they merit little further discussion.

26. The result for the elderly is fully consistent with the recent findings of McBean and Babish, 1995, on (raw) black-white differences in Medicare-reimbursed immunization rates.

Labor Market Behavior

Columns 2 and 3 display estimated negative *cet. par.* associations with borderline statistical significance between immunization propensities and WORKS or HOURS. A similar pattern is seen for the elderly sample in columns 4 and 5, although little attention will be paid henceforth to the elderly sample owing to its limited labor market participation. Were it assumed that the observed covariates X and the unobservables Θ are uncorrelated, then it might be tempting to infer from these point estimates that the "effect" of labor supply on immunization propensities is negative. Do workers perceive their time as so valuable that they tend to pass up immunization opportunities, for example? If so, then how would one structure cost-effective programs to encourage workers to obtain immunizations? Might a major component of such programs, designed accordingly, be structured to encourage workers to obtain flu shots by reducing the effective full prices of immunization they face? It turns out that the answers are not quite so simple. The analysis in the next section considers in greater detail the relationships between labor market behavior and immunization propensities, with the major focus there being whether the "ounce of prevention" or the "pound of cure" effect tends to be dominant with regard to labor market behavior.

The idea that unmeasured aspects of health status are one of the main elements of the omitted covariates Θ correlated with labor market behavior gains some support when one considers a specification comparable to that in column 2 of table 4 in which the self-reported health status measures are omitted. In this case, the point estimate (t-stat.) on WORKS becomes $-.018$ (3.7) as contrasted with the point estimate $-.007$ (1.6) reported in the table. Health status is clearly likely to be an important

structural determinant of both flu shot propensities and labor market behaviors, and it is clearly reasonable to expect that the set of self-reported health status measured used here is insufficiently rich to pick up all the important dimensions of such a covariance structure. However, as suggested earlier, the components of Θ correlated with labor market behavior probably consist of more than just unmeasured aspects of health status: time preference, risk aversion, subjective valuation of healthiness, unmeasured demographics, etc. are all likely to be important to some degree in influencing both labor market behavior as well as propensities to obtain immunizations.²⁷

VI. FLU SHOTS AND LABOR MARKET BEHAVIOR: A CLOSER INSPECTION

Table 5 presents the full set of reduced form and IV estimates of the various labor market effects in the non-elderly sample. Column 1 reports the results for the reduced form in which the state-level unemployment rate -- the baseline identifying instrument for WORKS and HOURS -- appears as the labor market measure. The estimates from two alternative IV strategies

27. To see that the unobservables consist to some degree of such components, a binary "use tobacco" variable was created for the non-elderly sample. Unobserved individual characteristics like time preference, risk aversion, valuation of health, etc. might in general be expected to have opposite impacts on tobacco use and immunization propensities, e.g. with all else equal, higher rates of time preference might work against immunization propensity but in favor of tobacco use.

When the tobacco use covariate is included in the specification comparable to that reported in column 1 of table 4, the resulting point estimate is $-.0123$ ($t=3.2$), thus providing some support for the notion that unobservables of the sort
(continued)

are reported in columns 3 and 4 for WORKS and columns 6 and 7 for hours. Finally column 8 reports the results of an OLS specification in which a predicted log-wage (LWAGE) measure is used as the labor market variable.²⁸ The OLS results from table 4 for WORKS and HOURS are reported again in columns 2 and 5 to facilitate comparisons.

The IV strategy (IV-1) adopted for the WORKS and HOURS specifications reported in columns 3 and 6, respectively, is a "just-identified" approach. The single labor market variable (WORKS or HOURS) is instrumented with the state-level unemployment rate (UE), the maintained assumption being that local labor market conditions structurally determine the WORKS and HOURS outcomes, but are uncorrelated with individual-level unobserved propensities to obtain flu shots.²⁹

described above are probably key determinants of immunization propensities.

28. It is this LWAGE result that enables one to most directly obtain a sense of how the individual's value of time might affect his/her propensity to obtain flu shots: Do individuals with relatively high time values perceive the time costs of obtaining a flu shot to be prohibitively high? Or might such individuals anticipate that the value of time lost from work if ill is sufficiently large that it is worth incurring the time costs of obtaining immunizations?

29. The reduced form results for WORKS and HOURS suggest that UE is probably a reasonable identifying variable for the labor market variables. In the reduced-form equations for WORKS and HOURS, the t-statistics for UE are 6.1 and 5.7, respectively, in both cases corresponding to negative point estimates.

It may be, however, that the apparently strong correlation between local unemployment rates and labor market outcomes holds only over part of the sample. For instance, the labor market outcomes for certain occupations may be determined strongly by local economic conditions (e.g. electrical contractors,
(continued)

For both WORKS and HOURS, the IV-1 point estimates are positive and statistically significant, consistent with the hypothesis that the treatment effect of labor supply on immunization propensities is positive. However, these point estimates are also quite large, probably unrealistically so. This could, of course, be a matter of IV bias due to underpowered instruments, but as indicated above the t-statistics on the UE point estimate in the reduced forms for WORKS and HOURS are quite large, so this is less of a concern than it might otherwise be. There is, of course, no way to test overidentification here since the model is just-identified. The most troublesome features of these results are the extremely large IV point estimates for WORKS and HOURS. While the state-level unemployment rate would seem to be a reasonable instrument, it is clear that by itself it explains little of the variation in the WORKS and HOURS variables.³⁰

The IV-2 results for WORKS and HOURS, reported in columns 4 and 7 are based on the assumption that there are additional covariates included in the definition of X that might properly be

construction workers) whereas the labor market outcomes for other occupations (e.g. college professors, professional athletes) may have little to do with local economic conditions. Schooling may be one observable that segments the sample in such a manner. Indeed, estimation of separate reduced forms for the subsamples for SCHOOLING<16 and SCHOOLING≥16 corroborate this to some degree: the point estimate for the UE parameter in the SCHOOLING<16 subsample is -.023 (t=7.2), whereas the corresponding estimate in the SCHOOLING≥16 subsample is .005 (t=1.1). As such, the contribution of information from subsamples having relatively high schooling attainment to identification and estimation of the labor market parameters may be limited when only UE is used as an instrument.

30. The R-squares in regressions of WORKS and HOURS on UE are only .0018 and .0017, respectively.

excluded from X but included in Z. Most of the covariates in X are likely to be important structural determinants of labor supply. In particular, the literature would suggest that family structure (described by FAMILY SIZE and MARRIED) is likely to be an enormously important determinant of labor supply behavior for both males and females. Because these covariates do not appear to be highly significant determinants of immunization propensity in the non-elderly sample,³¹ the IV-2 strategy is based on the exclusion of these covariates from X. In particular, this specification omits the FAMILY SIZE and MARRIED variables from X but includes them in Z; also included in Z is an interaction between MALE and FAMILY SIZE, the idea being that family size is likely to have a significant differential impact on labor supply for males and females.³² It might be pointed out that this alternative IV

31. This is emphatically not the case for the elderly sample.

32. Bound et al., 1995, Staiger and Stock, 1994, and others have suggested that IV estimates may be seriously biased unless in the reduced form regressions the F-statistics on the set of excluded regressors are large. For these data, this does not appear to be a concern. For the IV-2 instrument set, the F-statistics are quite significant for both the WORKS and the HOURS reduced forms. The point estimates (t-statistics) and corresponding F-statistics for the excluded variables from these reduced forms are:

Parameter	WORKS	HOURS

UE	-.016 (6.0)	-.77 (5.6)
FAMILY SIZE	.012 (4.0)	.91 (5.8)
FEMALE*FAMILY SIZE	-.083 (22.8)	-4.1 (21.9)
MARRIED	-.014 (2.1)	-1.1 (3.3)
F(4,21081)	190.2	185.4
	(p<.0001)	(p<.0001)

strategy results in overidentified models; as such, overidentification can now be tested via the J-test.

The IV-2 point estimates for both WORKS and HOURS are positive, statistically significant,³³ and, importantly, of magnitudes far more reasonable than the previous results. The overidentification tests are in both cases comfortably small.³⁴

33. To obtain some sense of the reliability of inferences based on the standard covariance matrix estimators -- with particular focus on the labor market and the WIDESPREAD FLU parameters -- a modest bootstrap exercise was undertaken. Five hundred bootstrap sample replications were undertaken for the model where WORKS appears as the labor market measure and in which the IV-2 instruments are used. Recall that the estimates based on the standard covariance estimators are .0568 (s.e.=.0219, t=2.587) for WORKS and .00155 (s.e.=.000695, t=2.225) for WIDESPREAD FLU.

The respective means/medians of the point estimate marginal bootstrap distributions are .0553/.0562 for WORKS and .00157/.00153 for WIDESPREAD FLU, with respective standard errors are .0211 and .000689. The naive empirical 95% confidence intervals are (.0129,.0951) for WORKS and (.000206,.002978) for WIDESPREAD FLU, while the 95% confidence intervals based on the method suggested by Davidson and MacKinnon, 1993 (which amounts to finding the shortest interval in the distribution of bootstrap point estimates containing 95% of the observations) are (.0129,.0947) for WORKS and (.000440,.00313) for WIDESPREAD FLU. All said, the bootstrap inferences are quite close to those based on the standard parametric methods. As such, key inferences based on the standard covariance estimators are likely -- at least in this sample -- to be quite reliable.

34. It has been suggested that J-tests for overidentification can have low power against many possibly interesting departures from the null. The results presented here, in which the null of orthogonality is virtually never rejected at standard confidence levels, might be subject to such a concern as well. To gain some sense of how well this test might pick up ostensibly important violations of orthogonality, a model where WORKS appears as the rhs labor market variable was specified, in which the family size and marital status variables were included in the covariate vector and the instrument vector. However, excluded from the covariate

(continued)

As such, the IV-2 results will be taken henceforth to be the "preferred" IV results for purposes of inference and interpretation.³⁵

vector are now UE and the four health status indicators (VERY GOOD, GOOD, FAIR, POOR). The thinking here is that self-perceived health status should influence both flu shot propensities as well as propensities to supply labor; as such, the exclusion of such measures should induce a correlation between unobservables and instruments. As above, it is maintained that UE identifies the model, so it is reasonable to test overidentification by means of the J-test.

In this case, the J-test χ^2 statistic turns out to be 20.8 on 4 d.f. ($p=.0003$), strongly suggesting a violation of orthogonality. As such, one might conclude that when at least some important forms of orthogonality violations are present in these data, the J-test is capable of detecting them reasonably well.

35. For purposes of comparison with the linear IV estimates, two alternative estimation strategies were considered.

First, probit specifications were estimated in which reduced forms for WORKS and HOURS were estimated as linear functions of the IV-2 instrument set. When the actual values of WORKS and HOURS are included as the probit regressors, the respective point estimates (t-statistics) are $-.025$ (0.8) and $-.00046$ (0.7). Replacing the actual values with the reduced form predictions gives point estimates of .49 and .010, respectively (the adjusted standard errors were not computed). These results are clearly qualitatively similar to those obtained using the linear model strategy.

Second, if instead of assuming that the distribution function $G(\cdot)$ is linear it is maintained that $G(\cdot)$ is exponential (i.e. $G(\cdot)=\exp(\cdot)$), an alternative IV estimation strategy can be implemented based on results from Mullahy, 1996. (Obviously this is at best an approximation since $\exp(\cdot)$ is not a valid distribution function. However, for outcomes -- like flu shots -- having relatively small conditional probabilities, the exponential assumption is likely to approximate reasonably well the shape of standard distribution functions like logistic and normal in the neighborhood of the relevant probabilities.) This exponential

(continued)

Finally, column 8 of table 5 reports the results of a linear probability model estimation where a predicted log-wage (LWAGE) is included as the labor market behavior measure.³⁶ The results for LWAGE are consistent with the IV results reported for WORKS and HOURS. That is, the positive and significant point estimate of the LWAGE parameter suggests that individuals having relatively high time values are those most likely to obtain immunization; income effects notwithstanding, these are also the people relatively most likely to be supplying hours to the labor market. As such, the results presented in columns 1, 3, and 5 of table 5 would suggest that the "ounce of prevention" effect is dominant.

Table 6 reports the results of a series of sensitivity analyses in which the sample is split alternatively by gender, race, schooling (0-12 vs. 13+ years), age (25-44 vs. 45-64), and marital status. The IV-2 definitions of X and Z are used in all these specifications.

model was estimated with the WORKS variable included as the labor market measure. The resulting point estimates (t-statistics) are -.063 (1.2) when WORKS is treated as exogenous versus .91 (2.7) when the IV estimation is based on the IV-2 instrument set.

On the basis of these sensitivity analyses, it is thus probably reasonable to suggest that the estimates of the labor supply effects presented in the text are not being unduly influenced by the linearity assumption.

36. Specifically, a standard Heckman-corrected log-wage prediction is obtained. The wage rate is computed as monthly earnings divided by monthly hours for observations where both are recorded, and is set to missing otherwise. The "instruments" used in the probit equation but omitted from the log-wage equation are the same as those specified in what will turn out to be the preferred IV-2 specification reported below: UE, FAMILY SIZE, MARRIED, and a MALE*FAMILY SIZE interaction. The particulars are available on request.

The WORKS and HOURS point estimates for males and females turn out to be reasonably close, although the differences are statistically significant. Moreover, the point estimates for males are relatively imprecise, possibly owing to the fact that the FAMILY SIZE and MARRIED instruments are stronger predictors of females' labor supply behavior than males'. For LWAGE the point estimate for females is again significant whereas that for males is not; in this case the magnitudes of the point estimates are themselves dramatically different as well.

For WORKS, HOURS, and LWAGE, the point estimates for nonwhites are all markedly smaller than those for whites, with the differences in all instances statistically significant. The relative imprecision of the point estimates for nonwhites owes in some measure to the smaller sample size.

The results for SCHOOLING are perhaps most interesting. No statistically significant differences in the point estimates for WORKS, HOURS, or LWAGE are apparent despite the fact that the individual point estimates are themselves (for the most part) significantly different from zero. The earlier speculation that UE might not be a particularly powerful instrument for labor supply in a relatively highly-educated population seems to have little implication for these estimates.³⁷

37. Not presented in these tables are the point estimates for SCHOOLING in the IV-2 specifications for WORKS and HOURS. It turns out that SCHOOLING still has a statistically significant structural impact on immunization propensities in the IV-2 specifications, albeit marginally smaller than that evident in the corresponding OLS specifications: .0017 (t=2.1) IV-2 vs. .0025 (t=3.3) OLS for WORKS, and .0015 (t=1.8) IV-2 vs. .0025 (t=3.3) OLS for HOURS.

When the sample is split into "young" (25-44) and "mature" (45-64) subsamples, it is somewhat curious that the point estimates of the labor market effects are for the most part statistically insignificant (the exception being LWAGE in the mature sample). Nonetheless, the magnitudes of the point estimates are markedly larger for the mature sample than for the young sample, and the differences are in all instances statistically significant. The statistical insignificance of these point estimates³⁸ might be attributable in part to important shared nonlinear age effects on both labor supply behavior and immunization propensity. It should also be noted that the only instance where the J-tests for overidentification would recommend rejecting the null of error-instrument orthogonality are for the ages 45-64 sample. Recalling that the point estimates of the MARRIED parameter are highly significant in the baseline specifications for the elderly sample (columns 4-6 in table 4), it is worth speculating that there may be a gradual increase with advancing age in the structural importance of such household characteristics whose onset begins well in advance of age 65; to the extent this is true, the exclusion restriction on MARRIED in X may be tenuous.

Finally, when the sample is split on the basis of marital status (presently married vs. presently unmarried), it is noteworthy that the point estimates for all three labor market measures are somewhat larger (though not statistically significantly so) for the unmarried subsample.

38. Or, perhaps more pessimistically, the statistical *significance* of the age-pooled results.

VII. SUMMARY

This analysis has offered some insights into the determinants of adult influenza immunization. The expected relationships between immunization status and both age and health status were largely apparent in the data, as were other results that are somewhat more novel.

First, a caveat. When interpreting these results, it should be remembered that the data are based on self- or proxy-reports of whether the individual received a flu shot in the past year and are, therefore, subject to recall and other reporting error. In general, factors like schooling attainment, age, the time of year at which the survey was administered, and whether the responses were provided by the individual or his/her proxy might be expected to influence the accuracy of self-reports of immunization status.

It would be quite useful to corroborate the "demand side" results obtained here using a "supply side" (i.e. clinical) database. To the extent that the inferences drawn here are valid, however, they should provide guidance to clinicians and policymakers as to how to enhance influenza immunization status should that be deemed a desirable objective.

The paper's main focus -- the structural relationship between labor supply and flu shot propensity -- has provided some interesting inferences whose utility may carry over to a broader consideration of preventive care issues. The fact that one often observes a *cet. par.* negative correlation in these data between the labor supply measures and flu shot propensity might lead one to conclude that individuals having relatively strong attachments to the labor market are inclined to ignore preventive care owing, e.g., to their relatively high time costs. The structural estimates presented here, in light of the theoretical framework

laid out in section III, indeed suggest that time costs may matter, but quite possibly in a very different way. That is, one interpretation of the IV results for the 25-64 sample is that individuals are more inclined to engage in preventive care precisely because of their relatively high time costs.

The results with respect to the previous flu season's severity -- the other novel innovation of this paper -- are of considerable interest. The WIDESPREAD FLU measure constructed here is admittedly a rather rough proxy for the abstract concept of "flu season severity," so some attenuation bias might be expected. Nonetheless, the results for both the pooled sample and for both age subsamples suggested a significant role for this phenomenon. One rational response to having suffered through, or having watched one's family, neighbors, or co-workers suffer through, bouts of flu one year -- presumably appreciating better than those not so afflicted and/or affected the costs of influenza -- is to engage in preventive activities before the next flu season arrives. Further exploration of this issue would clearly be useful.

In addition, the consistent findings on the relationship of immunization status to self-perceived healthiness have some particularly interesting policy implications. To the extent that individuals who perceive themselves to be healthy indeed have lower propensities to receive influenza immunizations (as the results suggest), then additional efforts may be needed to inform and convince such individuals -- particularly, though not exclusively, the elderly -- of the preventive benefits of immunization.

Indeed, the results for the nonelderly sample suggest that while workers may be less likely on average than nonworkers to

obtain flu shots, *working* is not the causal explanation. Instead, the results suggest that all-else-equal-healthier workers are likely to forego immunizations not because they are working but because they are healthier.

Discerning precisely why the healthy have, all else equal, lower propensities to be immunized is not possible with these data, but any number of possibly noncompeting explanations can be advanced. For instance, it might be the case that relatively less-healthy individuals -- who, presumably, have relatively more frequent contacts with their health care providers -- tend to receive their flu shots during provider contacts that are not sought primarily for immunizations. The fact that health care access measures (insurance and, in some cases, income) were estimated to be important determinants of immunization propensity corroborates to some degree such an access-based explanation.

Finally, the significant role estimated here for schooling may also be a policy-relevant finding. While better-educated individuals have repeatedly been shown to have greater demands for preventive health care, the line of causation determining such correlations has not been unambiguously established. Do the better-educated value healthiness more than the relatively less well-educated? Or do the better-educated better understand the medical benefits of various forms of preventive care (like flu shots)? Or is there some other explanation at work? To the extent that an accurate understanding of the medical benefits of flu shots is positively correlated with education, i.e. if information is a barrier in the status quo, then it *may* be the case that policies designed to better inform individuals about such benefits would result in enhanced immunization uptake.

Appendix: Demand Model Analytics

In addition to the health production function and utility function described in the main text, the model's constraints are as follows. The labor supply function is assumed to be given by

$$L = L(w, H) = (\lambda + \lambda_w w) \eta(H),$$

where w is the wage rate and where λ , λ_w , and $\eta'(\cdot)$ all positive.

λ can be viewed as the "rationed" or "fixed" level of labor supply, λ_w measures its responsiveness to wage rates, and $\eta(H)$ measures the impact on flu-related health outcomes on labor supply.³⁹ The labor supply function is specified in this manner so that the analytics undertaken here will be relatively comparable with the empirical analysis to follow, in which wage rate effects as well as labor supply effects are considered.⁴⁰ Leisure time is determined correspondingly as

39. To keep the analysis simple, any possible dependence of the wage rate on health outcomes is ignored. This would seem a reasonable restriction in the context of the short-run decision problem described here.

40. Health effects are specified as having proportional rather than additive effects on labor supply: It seems more reasonable to suggest that if having the flu reduces the labor supply of a 160-hour/month worker by, say, 24 hours (three work days), it would tend to reduce the labor supply of a 200-hour/month worker by 30 hours (three work days) rather than by 24 hours (2.5 work days). Note, too, that the no-rationing case is given by specifying $\lambda=0$ while the full-rationing case is given by specifying $\lambda_w=0$.

$$\ell = T - L(w,H) - \sigma S_\ell,$$

where T is total time available.

The amount of time required to obtain a flu shot is assumed to be a constant σ . If a shot is obtained by taking time off from work, then it is assumed that the individual loses $\alpha\sigma$, $\alpha \in [0,1]$, hours of paid work time, with $1-\alpha$ being thought of as the generosity of the employer's paid leave time. The out-of-pocket costs associated with obtaining the flu shot (office visit, vaccine, vaccine administration, etc.) are denoted $c_S \geq 0$. It is noteworthy that the "parameters" σ , α , and c_S can be manipulated in practice by a variety of policies and clinical practices. For instance, σ might be reduced by a policy that enhances employees' access to flu shots, e.g. workplace immunization campaigns. c_S is reduced in practice by public and/or private policies which call for the administration of flu shots at no or at reduced out-of-pocket cost.⁴¹

Finally, consumption is given by

$$Z = w(L(w,H) - \alpha\sigma S_W) - c_S S.$$

The analysis proceeds straightforwardly via comparisons of the three pairs of expected utility differences, $\Delta_{ij} = EU(S_i|\mathbf{X}) - EU(S_j|\mathbf{X})$ for $i,j \in \{W,\ell,N\}$ so that, in a stochastic context,

41. This has now been formalized in Medicare coverage policy, for example.

$$\begin{aligned}\text{Prob}(\text{Choose } i \text{ over } j | \mathbf{x}) &= \text{Prob}(V(S_i | \mathbf{x}) > V(S_j | \mathbf{x})) \\ &= G(\Delta_{ij})\end{aligned}$$

for some distribution function $G(\cdot)$ determined by the stochastic specification of $\eta_{ji} = \varepsilon_j - \varepsilon_i$. Comparative statics for covariate x_m are then given by $\partial \text{Prob}(\text{Choose } i \text{ over } j | \mathbf{x}) / \partial x_m = G'(\cdot) \partial \Delta_{ij} / \partial x_m$ with $G'(\cdot) > 0$. Since the data indicate only whether the individual obtained a flu shot -- not when -- the comparisons most relevant for purposes at hand are those for Δ_{WN} and $\Delta_{\ell N}$; for illustrative purposes, the analytics pertaining to Δ_{WN} are described here.

After some manipulation, Δ_{WN} can be written as

$$\begin{aligned}\Delta_{WN} &= (\pi_S - \pi_N) [\Psi(H_1) - \Psi(H_0) + \Lambda(T - \lambda\eta(H_1) - \lambda_w w \eta(H_1)) - \\ &\quad \Lambda(T - \lambda\eta(H_0) - \lambda_w w \eta(H_0)) + \\ &\quad (\eta(H_1) - \eta(H_0))(\lambda w + \lambda_w w^2)] - \beta - \alpha \sigma w - c_S \\ &= (\pi_S - \pi_N) [\dots] - \beta - \alpha \sigma w - c_S.\end{aligned}$$

Letting Λ_j be shorthand for $\Lambda(T - \lambda\eta(H_j) - \lambda_w w \eta(H_j))$, $j=0,1$, and recalling that $d \text{Prob}(\text{Choose } W \text{ over } N) / d(\cdot) = G'(\cdot) d\Delta_{WN} / d(\cdot)$, the main comparative statics results can be determined as follows:

$$\begin{aligned}d\Delta_{WN} / dw &= (\pi_S - \pi_N) [-\lambda_w (\Lambda_1' \eta(H_1) - \Lambda_0' \eta(H_0)) + \\ &\quad (\eta(H_1) - \eta(H_0))(\lambda + 2\lambda_w w)] - \alpha \sigma.\end{aligned}$$

$-\lambda_w (\Lambda_1' \eta(H_1) - \Lambda_0' \eta(H_0)) < 0$ from $\eta' > 0$ and $\Lambda'' < 0$, $-\alpha \sigma < 0$, and $(\eta(H_1) - \eta(H_0))(\lambda + 2\lambda_w w) > 0$, so the sign of the wage effect is

ambiguous. Thus, the dependence of Δ_{WN} of w depends on whether the value of leisure foregone from additional work effort plus the value of the time costs associated with obtaining the flu shot are larger or smaller than the value of additional consumption obtained from the additional labor supplied when healthy. By similar reasoning, the sign of

$$d\Delta_{WN}/d\lambda = (\pi_S - \pi_N) [-(\Lambda_1' \eta(H_1) - \Lambda_0' \eta(H_0)) + w(\eta(H_1) - \eta(H_0))]$$

is ambiguous. The signs of the remaining effects are unambiguous based on earlier assumptions; assuming [...]>0, then:

$$d\Delta_{WN}/dF = [...] \partial(\pi_S - \pi_N) / \partial F > 0$$

$$d\Delta_{WN}/d\alpha = -\sigma w < 0$$

$$d\Delta_{WN}/d\sigma = -\alpha w < 0$$

$$d\Delta_{WN}/dc_S = -1 < 0$$

$$d\Delta_{WN}/d\beta = -1 < 0$$

$$d\Delta_{WN}/dE = [...] \partial(\pi_S - \pi_N) / \partial E > 0.$$

These results indicate that the propensity to choose work-time immunization over non-immunization are positively related to the previous flu season's severity and to the anticipated exposure to infection, but negatively related to the out-of-pocket costs and the net time costs of immunization and to the degree of disutility associated with obtaining the shot itself.

References

- Advisory Committee on Immunization Practices (ACIP). 1994. "Prevention and Control of Influenza: Part I, Vaccines -- Recommendations of the Advisory Committee on Immunization Practices." *Morbidity and Mortality Weekly Report* 43, No. RR-9: 1-13.
- Ahituv, A., V.J. Hotz, and T. Philipson. 1996. "The Responsiveness of the Demand for Condoms to the Local Prevalence of AIDS." *Journal of Human Resources* (forthcoming).
- American College of Physicians, Task Force on Adult Immunization. 1990. *Guide for Adult Immunization*, 2nd Edition. Philadelphia: American College of Physicians.
- Borjas, G.J. and G.T. Sueyoshi. 1994. "A Two-Stage Estimator for Probit Models with Structural Group Effects." *Journal of Econometrics* 64: 165-182.
- Bound, J. et al. 1995. "Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak." *Journal of the American Statistical Association* 90: 443-450.
- Browning, M. and C. Meghir. 1991. "The Effects of Male and Female Labor Supply on Commodity Demands." *Econometrica* 59: 925-951.
- Chapman, L.E., et al. 1992. "Influenza -- United States, 1989-90 and 1990-91 Seasons." *Morbidity and Mortality Weekly Report* 41, No. SS-3: 35-46.
- Chapman, L.E., et al. 1993. "Influenza -- United States, 1988-89." *Morbidity and Mortality Weekly Report* 42, No. SS-1: 9-22.
- Davidson, R. and J.G. MacKinnon. 1993. *Estimation and Inference in Econometrics*. New York: Oxford University Press.
- Douglas, Jr., R.G. 1990. "Prophylaxis and Treatment of Influenza." *NEJM* 322: 443-450.

- Gardner, P. and W. Schaffner. 1993. "Immunization of Adults." *NEJM* 328: 1252-1258.
- Gill, P.W. and A.M. Murphy. 1985. "Naturally Acquired Immunity to Influenza Type A: Lessons from Two Coexisting Subtypes." *Medical Journal of Australia* 142: 94-98.
- Govaert, T.M.E. et al. 1994. "The Efficacy of Influenza Vaccination in Elderly Individuals." *JAMA* 272: 1661-1665.
- Hansen, L.P. 1982. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50: 1029-1054.
- Heckman, J.J. and T.E. MaCurdy. 1985. "A Simultaneous Equations Linear Probability Model." *Canadian Journal of Economics* 18: 28-37.
- Hoskins, T.W. et al. 1979. "Assessment of Inactivated Influenza-A Vaccine after Three Outbreaks of Influenza A at Christ's Hospital." *Lancet* 1: 33-35.
- Marwick, C. 1994. "Elderly Urged to Get Medicare-Paid Flu Shots; Agencies Try for Improved Awareness, Accessibility." *JAMA* 272: 1640-1641.
- McBean, A.M. and J.D. Babish. 1995. "Race-Specific Differences in Influenza Vaccination Levels Among Medicare Beneficiaries -- United States, 1993." *Morbidity and Mortality Weekly Report* 44: 24-27.
- McBean, A.M. et al. 1993. "The Impact and Cost of Influenza in the Elderly." *Archives of Internal Medicine* 153: 2105-2111.
- Mullahy, J. 1996. "Instrumental Variable Estimation of Count Data Models." *Review of Economics and Statistics* (forthcoming).
- Nguyen-Van-Tam, J.S. and K.G. Nicholson. 1993. "Influenza Immunization: Vaccine Offer, Request and Uptake in High-Risk Patients during the 1991/2 Season." *Epidemiology and Infection* 111: 347-355.

- Nichol, K.L. et al. 1995. "The Effectiveness of Vaccination against Influenza in Healthy, Working Adults." *NEJM* 333: 889-893.
- Philipson, T. 1995. "Private Vaccination and Public Health: An Empirical Examination for US Measles." *Journal of Human Resources* (forthcoming).
- Schmitz, R. et al. 1993. "Final Results: Medicare Influenza Vaccine Demonstration -- Selected States, 1988-1992." *Morbidity and Mortality Weekly Report* 42, No. 31: 601-604.
- Staiger, D. and J.H. Stock. 1994. "Instrumental Variables Regression with Weak Instruments." NBER Technical Working Paper No. 151.
- U.S. Dept. of Health and Human Services, Health Care Financing Administration. 1993. *Medicare Pays for Flu Shots* (pamphlet). Publication No. HCFA-10963.
- U.S. Food and Drug Administration. 1993. "New Drug for Flu." *FDA Consumer* (December): 2.
- U.S. National Center for Health Statistics. 1989. *Design and Estimation for the National Health Interview Survey, 1985-94*. Publication No. (PHS) 89-1384.
- U.S. Office of Technology Assessment. 1981. *Cost Effectiveness of Influenza Vaccination*. Washington: Office of Technology Assessment. Publication No. OTA-8-152.
- White, H. 1982. "Instrumental Variables Regression with Independent Observations." *Econometrica* 50: 483-500.
- Zamula, E. 1994. "How to Avoid the Flu." *FDA Consumer* 28: 16-19.

Figure 1
Flu Shot Decision Model

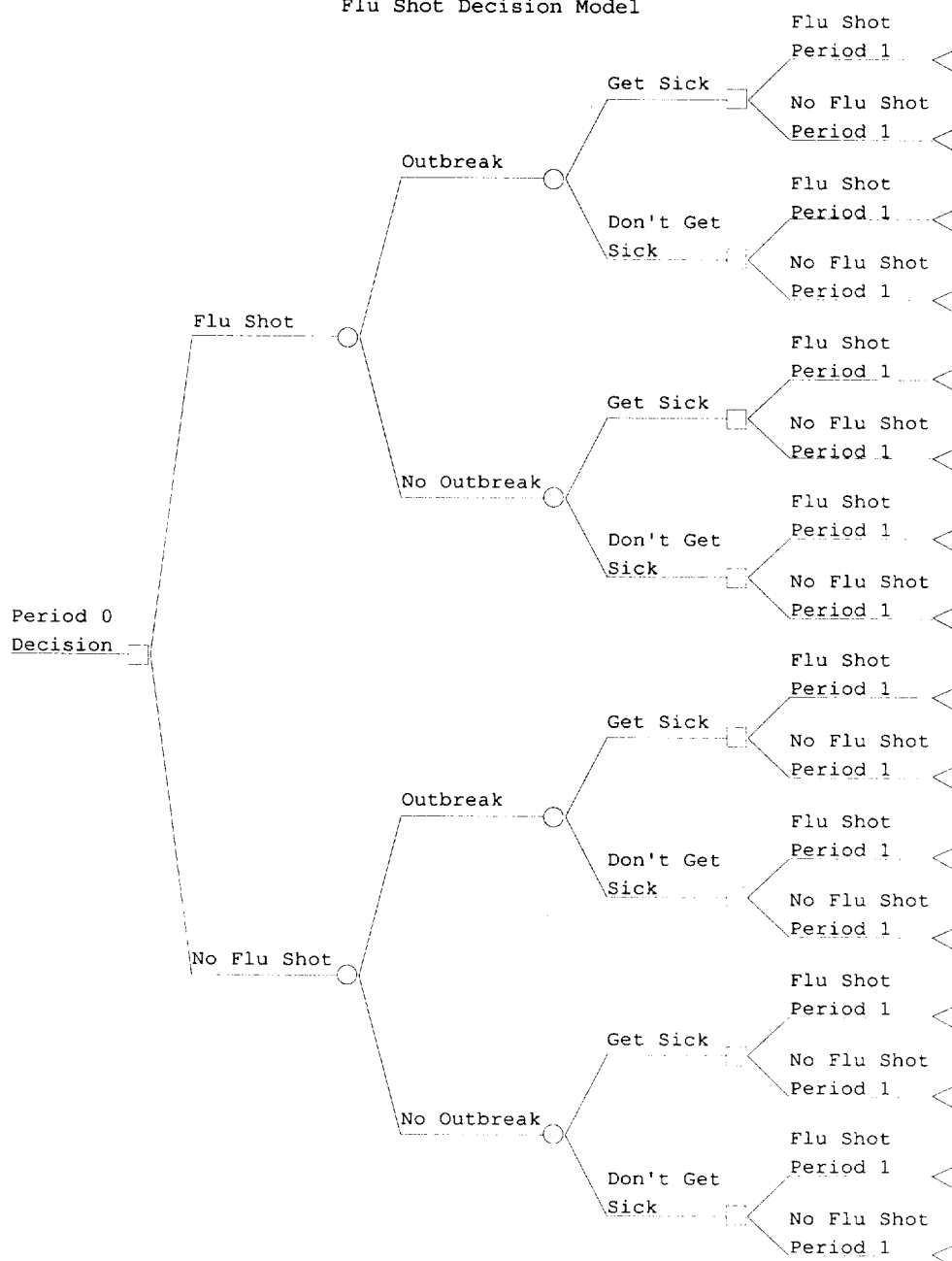


Figure 2

Flu Shot Propensity Age Profile

Jagged: Sample Mean by Age

Smooth: Predictions from Quartic Regression

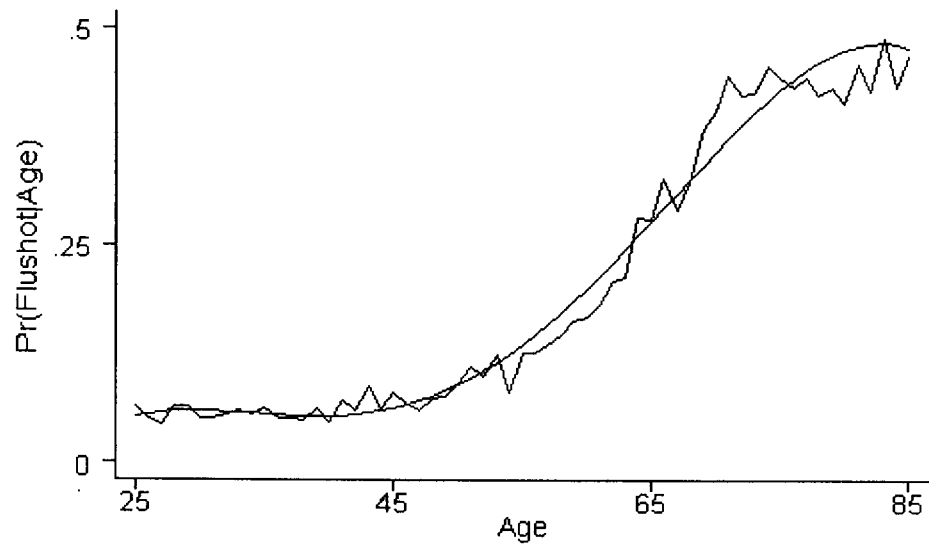


Table 1

Variable Definitions ("d.v." denotes 0-1 dummy variable)

Dependent Variable

FLU SHOT "During the past 12 months, have you had a flu shot?" (d.v.)

Labor Market

WORKS Individual reports usual activity is working (d.v.)
 HOURS Number of hours usually work per week
 LWAGE Log of monthly earnings divided by 4.33*HOURS

Risk Factors

WIDESPREAD FLU Number of weeks state reported widespread flu activity during
 1989-90 flu season
 HEALTHCARE At least one adult in household works in a health care
 WORKER occupation (d.v.)
 VERY GOOD Self-reported health status very good (d.v.)
 GOOD Self-reported health status good (d.v.)
 FAIR Self-reported health status fair (d.v.)
 POOR Self-reported health status poor (d.v.)
 AGE Age in years

Other Demographics

MALE Gender male (d.v.)
 SCHOOLING Number of years of completed schooling
 NONWHITE Race is not white (d.v.)
 INSURED Individual insured by some form of health insurance (d.v.)
 FAMILY SIZE Number of individuals in family
 MARRIED Individual is presently married (d.v.)
 NORTHEAST Individual resides in northeast census region (d.v.)
 MIDWEST Individual resides in midwest census region (d.v.)
 SOUTH Individual resides in south census region (d.v.)
 CENTER CITY Individual resides in central city part of an MSA (d.v.)
 OTHER MSA Individual resides in MSA but outside of central city (d.v.)
 UE 1991 average state-level unemployment rate

Table 2

Sample Means

	Ages 25-64 (N=21,103)	Ages 65+ (N=6,032)
<hr/>		
<i>Dependent Variable</i>		
FLU SHOT	.083	.400
<i>Labor Market</i>		
WORKS	.735	.100
HOURS	31.0	3.22
LWAGE	1.67	n.c.
<i>Risk Factors</i>		
WIDESPREAD FLU	2.67	2.70
HEALTHCARE WORKER	.044	.005
VERY GOOD	.294	.242
GOOD	.239	.320
FAIR	.077	.186
POOR	.030	.082
AGE	41.4	74.0
<i>Other Demographics</i>		
MALE	.432	.345
SCHOOLING	13.0	11.0
NONWHITE	.172	.124
INSURED	.863	.997
FAMILY SIZE	2.74	1.59
MARRIED	.628	.414
NORTHEAST	.225	.239
MIDWEST	.250	.265
SOUTH	.321	.336
CENTER CITY	.325	.330
OTHER MSA	.459	.386
UE	6.72	6.75

Table 3

Sample and Subsample Percentages: FLU SHOT
(N. obs. in parentheses)

	All Ages	Ages 25-64	Ages 65+
Both Sexes	.154 (27135)	.083 (21103)	.400 (6032)
Males	.141 (11203)	.081 (9119)	.403 (2084)
Females	.163 (15932)	.085 (11984)	.398 (3948)

Table 4

Baseline OLS Results
(Heteroskedasticity-consistent t-stats. in parentheses)

Variable	Ages 25-64 (N=21,103)			Ages 65+ (N=6,032)		
	1	2	3	4	5	6
Labor Market						
WORKS	--	-.007 (1.6)	--	--	-.015 (0.7)	--
HOURS	--	--	-.0001 (1.3)	--	--	-.001 (2.2)
Risk Factors						
WIDESPREAD FLU	.002 (2.4)	.002 (2.4)	.002 (2.4)	.008 (3.8)	.008 (3.8)	.009 (3.8)
HEALTHCARE WORKER	.060 (5.4)	.061 (5.4)	.061 (5.4)	.120 (1.3)	.126 (1.4)	.134 (1.5)
VERY GOOD	.007 (1.7)	.007 (1.7)	.007 (1.7)	.073 (3.8)	.072 (3.8)	.072 (3.7)
GOOD	.019 (3.9)	.019 (3.8)	.019 (3.8)	.095 (5.2)	.094 (5.2)	.093 (5.1)
FAIR	.055 (6.0)	.053 (5.8)	.054 (5.9)	.108 (5.2)	.106 (5.1)	.103 (4.9)
POOR	.102 (6.3)	.098 (6.0)	.099 (6.1)	.169 (6.2)	.167 (6.2)	.164 (6.0)
AGE	-.015 (9.5)	-.014 (9.2)	-.015 (9.3)	.098 (5.4)	.097 (5.3)	.095 (5.2)
AGE ²	.0002 (11.0)	.0002 (1.7)	.0002 (1.8)	-.001 (5.0)	-.001 (4.9)	-.001 (4.8)
Other Demographics						
MALE	-.001 (0.3)	.0002 (.04)	.0001 (.02)	-.009 (0.6)	-.008 (0.6)	-.005 (0.4)
SCHOOLING	.002 (2.9)	.002 (3.0)	.002 (3.0)	.015 (7.8)	.015 (7.8)	.015 (7.8)

(continued)

Table 4 (continued)

NONWHITE	.006 (1.0)	.006 (1.0)	.006 (1.0)	-.111 (5.8)	-.112 (5.8)	-.112 (5.8)
INSURED	.033 (6.9)	.034 (7.0)	.034 (7.0)	.130 (1.5)	.129 (1.5)	.131 (1.5)
FAMILY SIZE	-.002 (1.7)	-.003 (1.8)	-.003 (1.8)	-.013 (1.5)	-.013 (1.5)	-.013 (1.5)
MARRIED	-.0001 (.001)	-.00003 (.007)	-.0001 (.02)	.075 (4.6)	.075 (4.6)	.074 (4.5)
NORTHEAST	-.025 (4.2)	-.025 (4.2)	-.025 (4.2)	-.059 (2.8)	-.059 (2.8)	-.058 (2.8)
MIDWEST	-.025 (4.4)	-.025 (4.4)	-.025 (4.4)	-.026 (1.3)	-.026 (1.3)	-.026 (1.3)
SOUTH	-.006 (1.0)	-.006 (1.0)	-.006 (1.0)	-.016 (0.8)	-.016 (0.8)	-.016 (0.8)
CENTER CITY	-.006 (1.0)	-.006 (1.0)	-.006 (1.1)	-.001 (0.0)	-.001 (0.0)	-.001 (0.0)
OTHER MSA	-.020 (3.9)	-.020 (3.9)	-.020 (3.9)	-.025 (1.6)	-.025 (1.6)	-.025 (1.6)
CONSTANT	.273 (8.4)	.270 (8.4)	.270 (8.3)	-3.944 (5.7)	-3.882 (5.5)	-3.780 (5.4)

Table 5

Reduced Form and Structural Estimates: WORKS, HOURS, LWAGE (Ages 25-64)
 (Heteroskedasticity-consistent t-stats. in parentheses; LWAGE t-stats. not corrected for prediction;
 OLS estimates from Table 4 presented for comparison)

Variable	Reduced Form	WORKS			HOURS			LWAGE	
		OLS	IV-1	IV-2	OLS	IV-1	IV-2	OLS	OLS
Labor Market									
WORKS	--	-.007 (1.6)	.237 (2.0)	.057 (2.6)	--	--	--	--	--
HOURS	--	--	--	--	-.0001 (1.3)	.005 (2.0)	.0012 (2.6)	--	--
LWAGE	--	--	--	--	--	--	--	.041 (2.3)	--
UE	-.004 (2.1)	--	--	--	--	--	--	--	--
Risk Factors									
WIDESPREAD FLU	.001 (2.0)	.002 (2.4)	.001 (1.7)	.002 (2.2)	.002 (2.4)	.001 (1.9)	.002 (2.3)	.001 (1.9)	
HEALTHCARE WORKER	.061 (5.4)	.061 (5.4)	.041 (2.8)	.055 (4.8)	.061 (5.4)	.044 (3.2)	.055 (4.9)	.050 (4.2)	
VERY GOOD	.007 (1.7)	.007 (1.7)	.009 (2.0)	.008 (1.9)	.007 (1.7)	.012 (2.4)	.009 (2.0)	.007 (1.7)	

(continued)

Table 5 (continued)

GOOD	.019 (3.9)	.019 (3.8)	.029 (4.1)	.022 (4.2)	.019 (3.8)	.034 (3.7)	.023 (4.4)	.022 (4.3)
FAIR	.055 (6.0)	.053 (5.8)	.108 (3.8)	.068 (6.5)	.054 (5.9)	.109 (3.8)	.068 (6.5)	.067 (6.4)
POOR	.102 (6.3)	.098 (6.0)	.225 (3.5)	.131 (6.5)	.099 (6.1)	.215 (3.6)	.129 (6.6)	.129 (6.4)
AGE	-.015 (9.5)	-.014 (9.2)	-.027 (4.3)	-.019 (10.0)	-.015 (9.3)	-.027 (4.3)	-.018 (9.9)	-.018 (9.6)
AGE ²	.0002 (11.0)	.0002 (1.7)	.0004 (4.6)	.0002 (11.3)	.0002 (1.8)	.0004 (4.5)	.0002 (11.2)	.0002 (10.9)
<i>Other</i>								
<i>Demographics</i>								
MALE	-.001 (0.3)	.0002 (.04)	-.048 (2.0)	-.013 (2.0)	.0001 (.02)	-.059 (2.0)	-.016 (2.1)	-.018 (2.0)
SCHOOLING	.002 (2.8)	.002 (3.0)	-.00004 (.03)	.002 (2.1)	.002 (3.0)	-.001 (0.6)	.002 (1.8)	-.001 (0.7)
NONWHITE	.006 (1.1)	.006 (1.0)	.006 (1.1)	.006 (1.1)	.006 (1.0)	.011 (1.7)	.007 (1.3)	.011 (1.8)
INSURED	.033 (6.8)	.034 (7.0)	.019 (2.1)	.030 (5.9)	.034 (7.0)	.009 (0.7)	.027 (5.2)	.010 (0.9)
FAMILY SIZE	-.002 (1.7)	-.003 (1.8)	.006 (1.4)	--	-.003 (1.8)	.005 (1.3)	--	--
MARRIED	-.0001 (.01)	-.00003 (.007)	.0008 (0.2)	--	-.0001 (.02)	.003 (0.6)	--	--

(continued)

Table 5 (continued)

NORTHEAST	-.023 (3.9)	-.025 (4.2)	-.022 (3.3)	-.024 (4.0)	-.025 (4.2)	-.026 (4.1)	-.025 (4.2)	-.021 (3.3)
MIDWEST	-.027 (4.6)	-.025 (4.4)	-.024 (4.1)	-.025 (4.3)	-.025 (4.4)	-.028 (4.4)	-.025 (4.4)	-.020 (3.3)
SOUTH	-.006 (1.1)	-.006 (1.0)	-.011 (1.6)	-.007 (1.1)	-.006 (1.0)	-.014 (1.9)	-.007 (1.3)	-.001 (0.1)
CENTER CITY	-.005 (0.9)	-.006 (1.0)	-.006 (0.9)	-.006 (1.0)	-.006 (1.1)	-.003 (0.4)	-.005 (0.9)	-.014 (2.1)
OTHER MSA	-.020 (3.7)	-.020 (3.9)	-.024 (4.2)	-.022 (4.1)	-.020 (3.9)	-.021 (3.8)	-.021 (4.0)	-.032 (4.4)
CONSTANT	.299 (8.6)	.270 (8.4)	.354 (6.7)	.292 (8.8)	.270 (8.3)	.393 (5.6)	.302 (8.8)	.326 (8.1)
IV J-Test p-value	--	--	--	.40	--	--	.37	--

Table 6

Point Estimates for Selected Subsamples: WORKS, HOURS, LWAGE (Ages 25-64)
 (Results for WORKS and HOURS are from the IV-2 Specification, LWAGE is OLS;
 Heteroskedasticity-consistent t-stats. in parentheses;
 IV Overidentification Test p-values in square brackets;
 Wald Test p-values for Subsample Parameter Equality in curly brackets)

	Sex		Race		Schooling		Age		Marital Status	
	Male	Female	White	Nonwhite	≤12y	≥13y	25-44	45-64	Married	Unmar.
Full Sample (N=21103)	(9119)	(11984)	(17483)	(3620)	(11404)	(9699)	(13439)	(7664)	(13258)	(7845)
<hr/>										
WORKS										
.057	.061	.052	.056	.012	.067	.060	.028	.100	.049	.098
(2.6)	(0.7)	(2.3)	(2.7)	(0.2)	(2.1)	(2.0)	(1.3)	(1.6)	(1.4)	(2.0)
[.40]	[.82]	[.21]	[.10]	[.34]	[.42]	[.81]	[.19]	[.03]	[.43]	[.60]
	{.008}		{<.001}		{.187}		{<.001}		{.062}	
<hr/>										
HOURS										
.0012	.0010	.0011	.0012	.0003	.0015	.0012	.0006	.0014	.0013	.0021
(2.6)	(0.7)	(2.3)	(2.6)	(0.2)	(2.1)	(2.0)	(1.3)	(1.1)	(1.5)	(2.0)
[.38]	[.80]	[.22]	[.09]	[.34]	[.39]	[.78]	[.20]	[.02]	[.53]	[.51]
	{.006}		{<.001}		{.138}		{<.001}		{.068}	
<hr/>										
LWAGE										
.041	.001	.037	.047	.022	.039	.055	.021	.106	.036	.051
(2.3)	(.01)	(2.0)	(2.4)	(0.6)	(1.7)	(1.9)	(1.0)	(2.5)	(1.4)	(1.6)
	{.009}		{<.001}		{.214}		{<.001}		{.102}	