TEXT S1. ADDITIONAL METHODOLOGICAL DETAILS AND SENSITIVITY ANALYSIS

For

"Effect of the California Tobacco Control Program on Personal Health Care Expenditures"

by James Lightwood, Alexis Dinno, and Stanton Glantz

Tests for Stationarity

The Phillips-Perron unit root test was used to test for unit roots, using level or trend nonstationarity as the null [1]. (A unit root tests whether a variable y evolves as a random walk, i.e., with the dynamic relation $y_t = y_{t-1}$ + constant + deterministic trend + $error_t$.) The KPSS test [1,2], which uses level or trend stationarity as the null, was used to confirm the results of the Phillips-Perron test and to check for sensitivity or results to choice of null hypothesis. The Phillips-Peron test may have low power in small samples [1], so the KPSS test and analysis of ACF/PCAFs were used to stationarity when there was a conflict and Phillips-Perron showed borderline significant results. Unit root tests were done using a testwise 5% significance level using the year as defined for the raw data (calendar year for health care expenditures, fiscal year for the rest).

Automatic lag order selection was used for unit root tests and ACF/PACF analysis: ceiling($4[N/100]^{2/9}$) for Phillips-Perron; Hobijn [3] automatic bandwidth selection for KPSS test; min(ceiling [N/2-1], 40) for ACF/PACF.

The same procedures were used to test for stationarity in the regression residuals as for the dependent variables.

Regression Methods

We used standard time series regression techniques appropriate for non-stationary variables. Under the assumption of unit root, or random walk, nonstationarity, the appropriate procedure to determine any long run relationship between the variables is to estimate equations [1] and [2] as static regressions with no lagged variables. If these static regressions have stationary errors then they are called cointegrating regressions that describe the long run equilibrium relationship between the dependent and independent variables. Ordinary least squares estimates of cointegrating regression coefficients may poorly approximate t-distributions and be biased in small samples, so irrelevant instrumental variables (IIV) estimates [4,5] were used for estimation.

We use orthogonal basis function irrelevant instrumental variables (IIV) regression methods developed by Phillips [5] that are valid asymptotically with Monte Carlo evidence of good performance in small samples.

Adjustment for the long run covariance matrix of the Brownian motions of the regression variables (which is achieved by including first differences of explanatory variables in the IIV regressions) were not used for the estimates reported in Table 1. Inclusion of the first differences resulted in very small degrees of freedom and many apparent influential outliers. We believe that in this case omission of adjustment for the Brownian motions does not materially affect the conclusions. Ordinary least squares, and irrelevant instrumental variables regressions with and without adjustment for long run covariance of the random walks produced virtually identical results. Cointegrating regression slope coefficients converge at the rate of the number of observations (rather than at the square root for stationary variables), so stable estimates are expected even in small samples.

Sinusidual polynomial basis functions were used for IIV instruments as suggested in Phillips [5]; fourteen instruments were generated and identical instrument set was used for all IIV regression estimates. Sensitivity analysis indicated that results were insensitive to alternative choice for basis functions and number of instruments.

Lag order of 3 and 4 were used for conventional instrumental variables estimates assuming stationarity, due to evidence of moving average process at lags 2 and 3 in IIV estimates of equation [2].

The ECM regression coefficients converge at the rate of the square root of the number of observations, so small sample size is a concern when estimating the ECM. We used a 10% level of significance for the error correction term in order to estimate ECMs for the expenditure equation because of the low power in such a small sample (25 data points) for stationary variables. Lagged first differences of order one were included as needed to produce uncorrelated residuals. This model selection strategy was used because it was suspected that significance tests on the ECM equation coefficients had low power.

Parameter estimation was done using Stata Version 9 [6].

Results of Model Validity and Reliability Tests

1) In-sample Dynamic Predictions. The dynamic in-sample predictions of the cointegrating regressions for California per capita health care expenditure track the observed variables closely (Figure 1, Tables 1 and A1). The prediction error measures are similar to that of the cointegrating regression equation, and the regression results indicate that the actual observations are accurately predicted by the dynamic predictions. The short-run ECM model predictions perform slightly more poorly than those of the long run cointegrating regressions, particularly RMSE, for predictions conditional on observed and predicted (Figure 1) cigarette consumption, which may be due to the uncertainty because of stationary estimation with a small sample. The corresponding predictions for California per capita cigarette consumption (equations [1, 4]) and health expenditures also track the observed time series closely (Figure 1 and Table A1).

2) Sensitivity analysis of variation in control states and price indices. Sensitivity analysis using the regional and national MCPI, and varying the control states, did not change the results for equation [1] substantially (Table A2). Variation of price indices used to calculate real expenditures produced similar estimates that were virtually identical. Use of different states for control populations did not change the basic results: the coefficient for difference in cigarette consumption was always negative and different from zero. This coefficient varied from -15.8 (SE 1.83) when using Western region control states only to -33.8 (SE 2.24) when using all other states for control, which may be due to the effect of variation in non-smoking health related factors and their interaction with smoking in the

Table A1. Evaluation of dynamic in-sample model predictions

	Predicted Observations	Forecast errors		Regression, actual against forecast		
		Root mean square error	Absolute proportional error	R^2	Coefficient of simulated/forecast value	
					Estimate	SE
Health care expenditures conditional on observed cigarette consumption (2004 dollars per capita)						
Long-run*	24	44.9	0.00809	0.91	0.98	0.0637
Short run*	23	86.6	0.0163	0.71	0.81	0.114
Health care expenditures conditional on dynamic predictions of cigarette consumption (2004 dollars per capita)						
Long-run*	24	67.3	0.0114	0.91	0.98	0.064
Short-run*	23	111	0.0212	0.56	0.70	0.136
Difference in control and California per capita cigarette consumption (packs per capita)						
Long-run*	24	1.64	0.0373	0.98	0.99	0.029
Short-run	23	1.48	0.0302	0.98	0.97	0.027

Note: two observations are lost in simulation due to initialization of dependent variables for in-sample dynamic predictions.
*Long-run: predictions of the cointegrating regression describing the long-run relationship; short-run refers to adding. Short-run: predictions of the equilibrium correction model describing the short-run behavior of the variables.

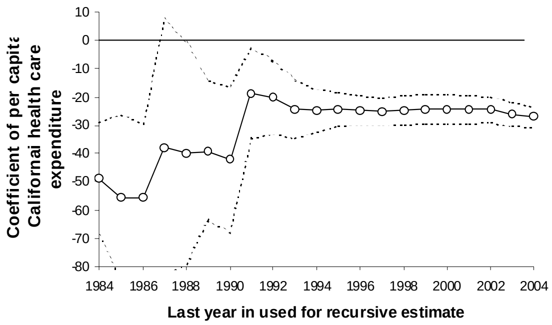

Figure A1. *Recursive estimates of coefficient for per capita California health expenditures in the cointegrating regression (equation [1]) show that the coefficient is very stable as annual observations are added to an initial estimate based on the years 1980 to 1984. The dotted lines are the 95% confidence interval for the estimates.*

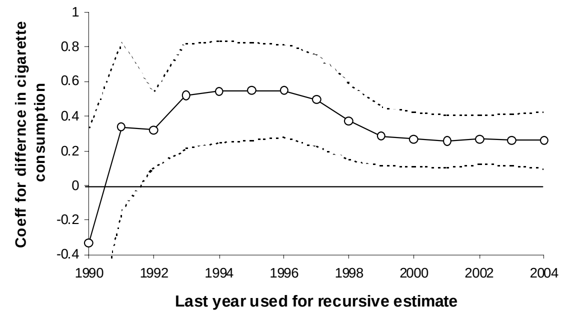

Figure A2. *Recursive estimates of coefficient for difference in cigarette consumption in cointegrating regression (equation [2]) are stable as more years of data are added to initial estimate computed based on the years 1980 to 1984. (There may be an structural break associated with cigarette tax increases that occurred between 1997 and 1999.) The dotted lines are the 95% confidence interval for the estimates.*

different control populations.

3*) Out of sample forecasts of the endogenous variables.* The out of sample performance of the cointegrating regressions for equation [1] were explored by estimating model for initial 14 (used only for expenditure only) and 16, 18 and 20 observations. These out of sample forecasts of California health expenditure (equation [1]) performed very well (Table A3), and were the result of extremely stable coefficient estimates over a range of the chosen forecast horizons, and within the 95% forecast intervals. The RMSE and mean absolute percentage error of the forecast regressions were also stable over the forecast horizon.

Corresponding forecasts for the difference in cigarette consumption (equation [2]) do not fare as well. Stable parameter estimates that produce accurate forecasts consistently within the forecast intervals (not shown) do not occur until estimation is performed over the first 20 observations. This poor forecasting performance may be due to structural break that occurred with large tax increases at the end of the 1990s. (See discussion of recursive estimates below.)

4) Reverse regression of parameter estimates for equations [1] and [2]. All the reverse regressions of the cointegrating relationships in both equations [1] and [2] returned similar renormalized coefficient estimates, as expected in cointegrating regressions (Table A4). The results on reverse regressions also are strong evidence that there is only one cointegrating relationship each in equation [1] and equation [2]. The invariance of the parameter estimates to the choice of normalization is important because it increases the likelihood that we are estimating a unique cointegrating regression connecting all the variables in the regression, rather than a composition of several cointegrating regression that may not have an interpretation consistent with the theoretical causal model for the regressions [7].

5) Models of bootleg cigarette sales. Time series models for untaxed (bootleg) sales were estimated. The best model used surveys of discount, Internet, Native American reservation, and Native American Internet real prices, which were between $1.60 and $2.00 per pack in year 2004 prices. Additional terms, functions of the difference in real prices of taxed cigarettes and an unweighted average of available survey data on discount cigarettes, was added to equation [1] to model the effect of possible bootlegging. These models of the effect of discount and bootleg cigarettes changed the cigarette consumption coefficient from -27.0 to -26.0. and were consistent with untaxed consumption

Table A2. Sensitivity analysis of total health care expenditures and cigarette consumption (equation [1])						
	Constant	Control expenditure	Difference in cigarette consumption	R^2	RMSE	Auto-correlaton
Original for total health care expenditures (from Table 1)						
Coefficient	2736	0.599	-27.0	0.91	46	0.09
SE	173	0.0519	1.82			
United States national MCPI deflator						
Coefficient	3267	0.519	-28.5	0.89	63	0.40*
SE	245	.0730	2.52			
Regional MCPI deflator						
Coefficient	3375	0.493	-28.2	0.94	46	0.03
SE	171	.0517	1.83			
All other states used as controls						
Coefficient	2606	0.615	-33.8	0.92	45	0.03
SE	186	.0528	2.24			
Northeast region controls only						
Coefficient	2867	0.458	-25.1	0.75	78	0.40*
SE	312	.0774	3.17			
Midwest region controls only						
Coefficient	2772	0.513	-21.4	0.90	49	-0.03
SE	180	.0482	1.53			
South region controls						
Coefficient	3346	0.588	-30.4	0.84	63	0.23
SE	193	.0730	2.97			
West region controls only*						
Coefficient	2563	0.523	-15.8	0.77	76	0.51*
SE	354	.0925	1.83			

*statistically significant auto-correlation in residuals at 5% significance level, Box-Ljung Q test
**cointegrating regression (residuals stationary) by Phillips-Perron unit root test at 10% level but not at 5% level. Residuals stationary by KPSS test at > 10% level, and inspection of ACF/PACFs.

Table A3. Evaluation of out-of-sample forecasts						
Sample periods used for estimate	Initial forecast year	Forecast errors		Regression, actual against forecast		
		Root mean square error	Absolute proportional error	R^2	Coefficient of simulated/forecast value	
					estimate	SE
Expenditure forecasts, cointegrating regression						
1994 to 2004	1994	69.0	0.0126	0.95	1.38	0.107
1996 to 2004	1996	71.8	0.0136	0.95	1.37	0.116
1998 to 2004	1998	87.5	0.0178	0.97	1.23	0.0906
2000 to 2004	2000	75.4	0.0135	0.97	1.33	0.124
Difference in cigarette consumption forecasts, cointegrating regression						
1996 to 2004	1996	11.0	0.212	0.045	0.0553	0.104
1998 to 2004	1998	11.2	0.231	0.064	-0.016	.00627
2000 to 2004	2000	3.63	0.064	0.41	-1.12	0.319

between 5% and 10% of total consumption in California.

6) Recursive estimates of coefficients. Recursive estimates of the difference in cigarette consumption equation [1] were extremely stable over the whole sample period (Figure A1)..

Recursive estimates revealed that the parameter estimates for the coefficient for the effect of cumulative California tobacco control program expenditures (β_1) in equation [2] are stable over successive sub-samples, though may experience a structural shift around 1997-1998. The recursive estimates for the cigarette price coefficients in equation [2] (Figure A2) indicate a structural break just before 2000, and most out of sample forecast inaccuracy is do to a sudden increase in sensitivity of cigarette consumption in both California and control states to price after 1998. These results show that the estimated models were quite stable over the observation period.

7) Exogenous time trends. Deterministic time trends did not chnage the cointegrating relationship for California health expenditures (equation [1]).
Alternative models using quadratic and cubic deterministic time trends in the cigarette demand equation (equation [2]) produced similar results, with somewhat larger effects of tobacco control expenditures in reducing cigarette consumption in California. A linear time trend only was used for the final estimates because it produced results that were less sensitive to outliers, and more stable recursive and rolling subsample estimates.

8) Unrestricted estimation of California health expenditures. Removing the restriction that the coefficients of California and control state cigarette consumption be equal and of opposite sign do not change the regression estimates for equation [1].

The estimates of the cointegrating estimates were robust to the remaining sensitivity analyses. Alternative methods of estimating equations [1] and [2] (ordinary least squares, instrumental variables estimates using lagged variables as instruments, Johansen maximum likelihood estimates, Prais-Winston autoregressive estimates) produce similar

Table A4. Reverse regressions for cointegrating relationships

California Health Expenditure $h_{CA,t}$ (equation [1])

Dependent variable for reverse regression estimate	constant	Control state health expenditure	Difference in cigarette consumption
Control health expenditures, $h_{control,t}$	2179 (219)	0.663 (0.0863)	-29.0 (1.99)
Difference in cigarette Consumption, $(s_{control,t} - s_{CA,t})$	2599 (194)	0.644 (0.0553)	-28.7 (0.0679)

Difference in cigarette consumption $(s_{control,t} - s_{CA,t})$ (equation [2])

Dependent variable for reverse regression estimate		CA tobacco control educational expenditure	CA cigarette price	Control state cigarette price	Time
CTCP cumulatve education expenditures, E_t	32.9 (2.65)	0.428 (0.262)	10.6 (2.49)	-22.4 (3.26)	1.39 (0.246)
California cigarette price, $p_{CA,t}$	30.8 (2.36)	0.246 (0.0847)	14.4 (0.187)	-26.2 (4.38)	1.69 (0.203)
Control cigarette price, $p_{control,t}$	31.4 (2.24)	0.260 (0.0800)	13.1 (0.0655)	-25.3 (2.043)	1.72 (0.192)
Elapsed time, $(t - t_0)$	30.1 (2.19)	0.214 (0.0814)	11.3 (2.24)	-23.0 (2.96)	1.84 (0.100)

The reverse regressions are estimated using variable listed in left-hand column as dependent variable, then renormalized so that coefficient California health care expenditure (top) or difference in cigarette consumption (bottom) is equal to one, in order to compare to estimates with normalization presented in Table 1.
Approximate standard errors for IIV reverse regression calculated by the delta method [8]

results. Estimators not using instrumental variables produced systematically different coefficient estimates for the cigarette demand equation (equation [2]). This was in contrast to the expenditure equation (equation [1]), where all estimators produced nearly identical results. This result is consistent with the endogeneity expected in the explanatory variables in equation [2], and the results of the ECM equation estimates. All of the explanatory variables in equation [1] are exogenous, while for equation [2] there are theoretical reasons and evidence from ECM equations indicating that some of the explanatory variables (price of cigarettes) are endogenous. The ECM equations did not indicate endogeneity of cigarette prices in equation [2]. Price changes appeared to be mostly due to several large and presumably exogenous tax increases over the sample period, and may explain this result. No formal Hausman type tests were done, however.

Other tests and estimates. Additional variables were included in equation [1] to measure differences between California and control states. Demographic differences were measured by proportion of resident population age 65 years or older in both populations. Differences in economic activity were measured by per real per capita income (year 2004 dollars) in both populations using data from the Bureau of Economic Analysis, deflated by the all-item consumer price index for urban consumers. Differences in health care market structure were measured by proportion of Medicare recipients enrolled managed care in both populations, and the ratios of licensed physicians, and licensed acute care community hospital beds to resident populations of California and control state populations. The measure of managed care market penetration closely parallels the proportion of the total population enrolled in managed care programs in California and control states. Market penetration was calculated using annual issues of the National HMO Census, Kaiser Family Foundation reports, annual issues of Medicare Program Statistics reports, and data from the Centers for Medicare and Medicaid Services. The ratio of licensed physicians and hospital beds to the resident populations were taken from various issues of the Statistical Abstract of the United States, and Health United States.

Adjustment for the differences in the proportion of elderly in the populations, per capita personal income, managed care market penetration, physicians and licensed hospital beds between California and control populations did not change the estimates for equation [1] significantly.

Only the coefficient for the age adjustment variable showed consistent borderline statistical significance with the expected sign (that is, costs increase with proportion age 65 or older). The information of these additional variables was summarized using a principal components analysis based on the covariance matrix of rescaled variables; the largest two principal components explained 99% of the total variation. Inclusion of these two principal components did not significantly change the estimates for equation [1] reported in Table 1.

A similar analysis was done using the prevalence behavioral health risk factors for California and control states taken from the Behavioral Risk Factor Surveillance Survey. Exploratory adjustment using the prevalence of overweight, obesity, binge drinking and hypertension among those with history of a blood pressure check did not significantly change the results. These risk factors with were also used in a principal components analysis and the two factors that accounted for over 95% of total variation in the data set. Neither component significantly affected the results. A final exploratory analysis was done by replacing the proportion of population that was elderly with a principal components analysis of the age structure of the population (age groups 0-5, 5-17, 18-24, 25-44, 45-64, 65+); this analysis also did not significantly change the results.

Comparison with Distributed Lag Models Assuming Stationarity

Exponential distributed lag models using annual California tobacco control program expenditures in the cigarette demand equations (equation [2]) produced results roughly consistent with the cointegrating regression. The estimated annual decay of the effect of annual tobacco control expenditure on the difference in cigarette consumption was statistically significant and between 0.12% and 0.05%, depending on the regression specification. These results indicate that annual tobacco control expenditures have persistent effects. The distributed lag model performed as well as the cointegrating model by standard in-sample statistical measures such as the F test for statistical significance and R^2. However the residuals displayed significant autocorrelation and dynamic predictions of the dependent variable did not perform as well as the cointegrating and ECM regressions.

Fiscal vs. Calendar Year Time Aggregatation. Using lagged unsmoothed fiscal year tobacco control program expenditures did not significantly change the results of any of the equations. Conversion of all variables to calendar year had no material effect on the results.

State Residential Personal Health Care Expenditures

The current analysis used all-payer expenditure data for each state. Corresponding data for payments to state residents are available from the Centers for Medicare and Medicaid Services only for the years 1991 to 1998. The resident data series are not long enough to re-estimate the model reliability, therefore, a correlation analysis was used to determine how well the all-payer data represented the resident data for the available data. The pairwise Pearson correlation coefficients and canonical correlations were used to measure agreement between the resident and all-payer expenditure series. If these correlations are high, an analysis with resident data (were it available) should confirm that using all-payer data. Because of the small sample size, the correlations should be considered summary statistics. The resident and all-payer data series differed by less than 1%. All correlations between the resident and all-payer series in levels and first differences were greater than 0.93. The canonical correlations between the all-payer and resident data were greater than 0.96 in levels and first differences. The available data indicate that the all-payer data can be used to represent the resident data for California and the 38 control state populations.

References

1. Maddala GS, Kim I-M (1998) Unit Roots, Cointegration, and Structural Change. Cambridge: Cambridge University Press.
2. Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y (1992) Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic times eries have a unit root? Journal of Econometrics 54: 159-178.
3. Hobijn B, Franses PH, Ooms M (1998) Generalizations of the KPSS-test for Stationarity. Rotterdam, Netherlands: Econometric Institute, Erasmus University. Econometric Institute Report, no. 9802/A Econometric Institute Report, no. 9802/A.
4. Phillips PCB, Hansen BE (1990) Statistical Inference in Instrumental Variables Regression with I (1) Processes. The Review of Economic Studies 57: 99-125.
5. Phillips PCB (2006) Optimal estimation of cointegrated systems with irrelevant instruments. Cowles Foundation. Yale University.
6. StataCorp LP (2003) Stata: College Station, Texas.
7. Enders W (2004) Applied Econometric Time Series. Hoboken, NJ: John Wiley and Sons.
9. Greene WH. (2000) Econometric Analysis, 4th ed. Upper Saddle River, HJ. Prentice-Hall.