Parental Substance Use and Foster Care:
Evidence from Two Methamphetamine Supply Shocks

Scott Cunningham and Keith Finlay

Supplementary Materials

Cleaning STRIDE data to generate market prices

We largely follow the methodology that Arkes et al. (2004) outline to prepare a series of meth prices. This report, which the authors prepare for the White House Office of National Drug Control Policy, examines the price trends for cocaine, heroin, cannabis, and meth in the US using prices from the Drug Enforcement Agency's System to Retrieve Information from Drug Evidence (STRIDE). We acquired STRIDE through a Freedom of Information Act request. STRIDE observations come from law enforcement events such as lab seizures, undercover purchases, etc. Samples are sent to DEA labs to identify the drugs and purities. Cocaine, heroin, and meth, occur sufficiently frequently to construct a price series. On the other hand, law enforcement officers collect most cannabis observations from seizures rather than purchases, and therefore it is not possible to construct a marijuana price series.

Following Arkes et al. (2004), we keep US observations originating from undercover purchases, individual seizures, and lab seizures and drop observations with missing or nonsensical price, weight, or purity data. We link drug observations to a drug market analogous to a metropolitan statistical area. Observations outside of major metropolitan drug markets are assigned markets associated with Census divisions.

Each observation is assigned a market quantity or distribution level based on net weight from the sample. For meth, we use three market quantities defined as having a net weight of less

than ten grams, between ten and 100 grams, and more than 100 grams. In this paper, we call meth observations retail if they come from the smallest two categories (i.e., less than 100 grams).

With the samples and market quantities defined, prices are regression adjusted to account for variation in sample purity. These regression models incorporate drug market random effects according to the following model:

$$purity_{ijk} = \alpha_{0k} + \alpha_{1k}time_{ij} + \alpha_{2k}weight_{ijk} + \varepsilon_{ijk},$$

where $time_{ij}$ is a vector of dummy variables representing a year-month and $weight_{ijk}$ is the raw weight of the ith observation in city k at time j. The coefficient, α_{0k} represents the intercept for city k, α_{1k} is a vector for the time coefficient for city k, and α_{2k} is the amount coefficient for city k. The disturbance term ε_{ijk} is distributed iid from normal distribution with mean zero. Our model is a random coefficients model where:

$$\alpha_{0k} = \gamma_0 + u_{0k},$$

$$\alpha_{1k} = \gamma_1 + u_{1k}, \text{ and}$$

$$\alpha_{2k} = \gamma_2 + u_{2k},$$

where γ_0, γ_1, and γ_2 are, respectively, the overall mean estimates for the intercept, time, and amount effects. The random coefficients for the intercept, amount and time are each assumed to be iid across cities and distributed

$$\begin{pmatrix} u_{0k} \\ u_{1k} \\ u_{2k} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \begin{vmatrix} \tau_{00} & 0 & 0 \\ 0 & \tau_{11} & 0 \\ 0 & 0 & \tau_{22} \end{vmatrix} \right).$$

Unlike Arkes et al. (2004), our specification uses month-year for time instead of quarter-year. We also constrain the off-diagonal elements of the random coefficient variance-covariance matrix at zero. This was done for computational reasons, as our models would not otherwise

2

converge. This accounts for the within-city clustering of the intercept, time and amount, but requires that across-city correlations be zero.

After estimating the purity equation, we retain the fitted values to predict purity ("*purity*"), which is then used to estimate the following price equation:

$$E(\textit{real price}_{ijk} \mid \gamma_{0k}, \gamma_{1k}, \gamma_{2k}) = \exp(\gamma_{0k} + \gamma_{1k}\textit{time}_i + \gamma_{2k}[\ln(\textit{weight}_{ijk})+\ln(\textit{purity}_{ijk})])$$

$$\gamma_{0k} = \beta_0 + c_{0k},$$

$$\gamma_{1k} = \beta_1 + c_{1k}, \text{ and}$$

$$\gamma_{2k} = \beta_2 + c_{2k},$$

$$\begin{pmatrix} c_{0k} \\ c_{1k} \\ c_{2k} \end{pmatrix} \sim N \left(\begin{vmatrix} 0 \\ 0 \\ 0 \end{vmatrix} \begin{vmatrix} \tau_{00} & 0 & 0 \\ 0 & \tau_{11} & 0 \\ 0 & 0 & \tau_{22} \end{vmatrix} \right).$$

The real price for observation i in period j in city k is modeled as a function of time, city effects, and the sum of the natural logarithm of amount and the natural logarithm of expected purity estimated in the previous regression. The mean effects of the control variables' effect on price are captured in the estimated β terms. The γ_0, γ_1, and γ_2 coefficients are assumed to be drawn from a normal distribution with mean zero.

We estimate the model using a linear mixed model. Except for our modeling of time as month-year and the imposed additional structure that the off-diagonal elements of the variance-covariance matrix be zero, our model is the same as that specified in Arkes et al. (2004).

Timing interventions and constructing the meth price instrument

To time the interventions, we use a stepwise regression procedure using the following model:

$$E(\textit{real price}_{ijk}) = \delta_0 + \tau_t + \tau_t^2 + \upsilon_{it},$$

where expected price is a variable of individual meth price observations, τ_t is a linear time trend common to all states, and τ^2_t is a quadratic time trend common to all states. We start without any fixed effects for the intervention months. Stepwise, we add a single fixed effect for each month after the intervention. If the fixed effect is significant, we keep it in the model. We continue these steps until a post-intervention, contiguous-month fixed effect is no longer significant. Using this procedure, we obtain the intervention lengths.

To estimate the meth price instrument, we estimate the following model:

$$\mathrm{E}(\textit{real price}_{ijk}) = \delta_0 + \tau_t + \tau^2_t + \phi_t \, \mathrm{I}[\textit{intervention}_t] + \upsilon_{it},$$

where expected price is a variable of individual meth price observations, τ_t is a linear time trend common to all states, τ^2_t is a quadratic time trend common to all states, ϕ_t is a month fixed effect, and I[intervention$_t$] is an indicator for months during supply interventions. In figure 3, we show the time series of the data as the ratio of median monthly expected retail prices for meth, heroin, and cocaine relative to their respective values in January 1995.

The price deviations, which form the instrumental variable used in the two-stage least squares modes, are defined as follows:

$$\textit{price deviation}_t = \phi_t \text{ during interventions, and 0 otherwise.}$$

Figure A1 shows the estimated quadratic time trends in prices as well as the price deviations estimated from the model.

AFCARS and TEDS data quality

We generate a number of data quality indicators to control the regression samples. We exclude Alaska, the District of Columbia, New Mexico, and South Dakota from regressions of all AFCARS outcomes. We drop New York from all route of admission into foster care regressions,

and Illinois from parental drug use regressions. These states have incomplete route information for the latest removal for the child in foster care.

We exclude Arizona, the District of Columbia, Kentucky, Mississippi, West Virginia, and Wyoming on the basis of poor TEDS data quality. These states either have poor data quality in general or for meth in particular.

The net result of these sample truncations is the removal of Arizona, the District of Columbia, Kentucky, Mississippi, New Mexico, South Dakota, and West Virginia from our regressions. Most of these states are small and tend to take longer to fully interface the AFCARS and TEDS federal data systems. Arizona, New Mexico, and South Dakota are the most substantive losses because those states had growing meth use during this period. The other states had much smaller meth user populations during this period.

Figure A2, Panel A, shows the proportion of the US population aged 0–19 years covered by our sample over time. Since our regressions use the same weights as this figure, it is clear that most of the identification from the models comes from the second supply intervention. The first supply shock does help identify the model for states with early AFCARS participation. Figure A2, Panel B is analogous to Panel A, but instead uses total state meth treatment admissions in the last year of the sample as the weight for each state. This figure shows that the states missing from the sample tend to come from the types of states with less meth use at the end of the sample.

Descriptive statistics from regression sample

Table A1 shows the descriptive statistics for monthly methamphetamine (meth) treatment and foster care admission flows that we use for our analysis. Our measurement of foster care

admissions and exits is from the Adoption and Foster Care Analysis and Reporting System (AFCARS). During the sample period, 221 white children entered and 37 exited foster care in an average state-month. Disaggregated by route, nine white children were placed in foster care due to parental incarceration, 93 due to parental neglect, 26 due to parental drug use, and 36 due to parental use. There were 37 exits from foster care in an average state-month during our sample period as well.

The Treatment Episode Data Set (TEDS) records information on every individual patient who received treatment for substance abuse from federally funded treatment facilities. As nearly all treatment facilities receive at least some federal funds, this constitutes a near census of the population of treatment admissions. We collect information on meth, alcohol, cocaine/crack, heroin and marijuana admissions based on whether any of the substances were mentioned in the patient's primary, secondary, or tertiary substance used at the last substance abuse episode prior to admission. During the sample period, 245 individuals in an average state-month were admitted for meth use, with 78 on average entering due to self referral. Alcohol was the most frequently mentioned drug in a patient's file (1,506 in an average state-month), followed by marijuana (712 mentions), cocaine/crack (525 mentions) and heroin (377 mentions).

Table A1 also provides information on control variables used in our models. The mean unemployment rate by state-month was 4.37%, and the mean cigarette tax per pack was $0.36. Unemployment statistics were collected from the Bureau of Labor Statistics, and cigarette tax data were collected from Orzechowski and Walker (2008). Population statistics are linear interpolations from the SEER data. The mean number of white 0 to 19 year olds for every one thousand persons by state-month was 1,491 and the corresponding statistic for white 15 to 49 year olds was 2,776. Table A1 also reports information about our instrumental variable, the

deviation in the real price of a pure gram of meth from its long run trends, measured at both the national and Census-division levels.

Assessing bias caused by endogeneity and measurement error

Here, we assess how measurement error and omitted variables bias may influence our OLS and IV estimates. Without loss of generality, let us ignore the panel aspect of the data and suppose the foster care model includes no covariates other than meth use and an unobservable factor W:

$$\log(\text{foster care}) = \alpha + \beta \log(\text{meth use}) + \gamma W + e.$$

If $\log(\text{meth use})$ is exogenous conditional on W, then β is the causal parameter of interest. Since we cannot observe W, there is endogeneity bias. In addition, since meth use is an illicit activity, we cannot observe the number of meth users, so we use meth treatment admissions as a proxy. Then, our estimating equation is:

$$\log(\text{foster care}) = \alpha_{OLS} + \beta_{OLS} \log(\text{meth treatment}) + u.$$

Suppose that a constant proportion, $0 < \zeta < 1$, of total meth users are in treatment at any given time with a multiplicative white noise measurement error, η:

$$\log(\text{meth treatment}) = \log(\zeta \text{ meth use}) + \log(\eta).$$

Substituting into the estimating equation, we have:

$$\log(\text{foster care}) = \alpha_{OLS} + \beta_{OLS} [\log(\zeta \text{ meth use}) + \log(\eta)] + u$$

$$= [\alpha_{OLS} + \beta_{OLS} \log(\zeta)] + \beta_{OLS} \log(\text{meth use}) + [u + \beta_{OLS} \log(\eta)].$$

Therefore, if the model is in logs and the assumption holds that the treatment population is some fixed proportion of the meth-using population (times an iid error), this measurement error resembles substantively that of classical errors-in-variables. The scale parameter ζ is absorbed

into the constant term and the proxy error is absorbed into the error term. Therefore, we can use two-stage least squares to estimate the causal parameter β.

References

Arkes, Jeremy, Rosalie Liccardo Pacula, Susan M. Paddock, Jonathan P. Caulkins, and Peter
Reuter. 2004. *Technical Report for the Price and Purity of Illicit Drugs through 2003*.
Santa Monica, CA: Rand.

Orzechowski and Walker. 2008. *The Tax Burden on Tobacco: Historical Compilation*. Volume
43. Arlington, VA: Orzechowski and Walker.

Table A1: State-varying variables selected descriptive statistics, whites (for AFCARS and TEDS variables), 1995–1999

| Variables	Source	Obs.	Mean	S.D.	Min.	Max.
Foster care at latest date of entry	AFCARS	1,428	221	237	6	1,356
-by parental incarceration		1,404	9	12	0	67
-by parental neglect		1,404	93	113	0	783
-by parental drug use		1,344	26	40	0	376
-by parental abuse		1,404	36	38	0	284
Foster care exits		1,428	37	43	0	277
Meth admissions	TEDS	1,428	245	591	0	3,638
-by self-referral route		1,428	78	218	0	1,505
Alcohol admissions		1,428	1,506	1,489	23	10,253
Cocaine/crack admissions		1,428	525	529	4	4,155
Heroin admissions		1,428	377	675	0	3,426
Marijuana admissions		1,428	712	671	11	3,935
Unemployment rate	BLS	1,428	4.37	1.17	1.7	8.5
Cigarette tax per pack (2002 $)	Orzechowski and Walker (2008)	1,428	0.36	0.21	0.02	0.93
Population 0-19 year olds (1,000s)	SEER	1,428	1,491	1,651	88	8,026
Population 15-49 year olds (1,000s)		1,428	2,776	2,922	198	13,953
Real price of meth deviations	STRIDE	1,428	52	82	0	416
-by Census division		1,428	53	110	-69	1,264

Notes: All variables are measured at the month-state level. AFCARS variables measure entry and exit for white children only. TEDS variables measure admissions for whites only. The number of observations for latest date of entry into foster care and route of latest entry can differ because not all states report all routes of entry.

Figure A1:   Density of state meth price observations (minimum, median, and maximum) by month, STRIDE, 1995−1999
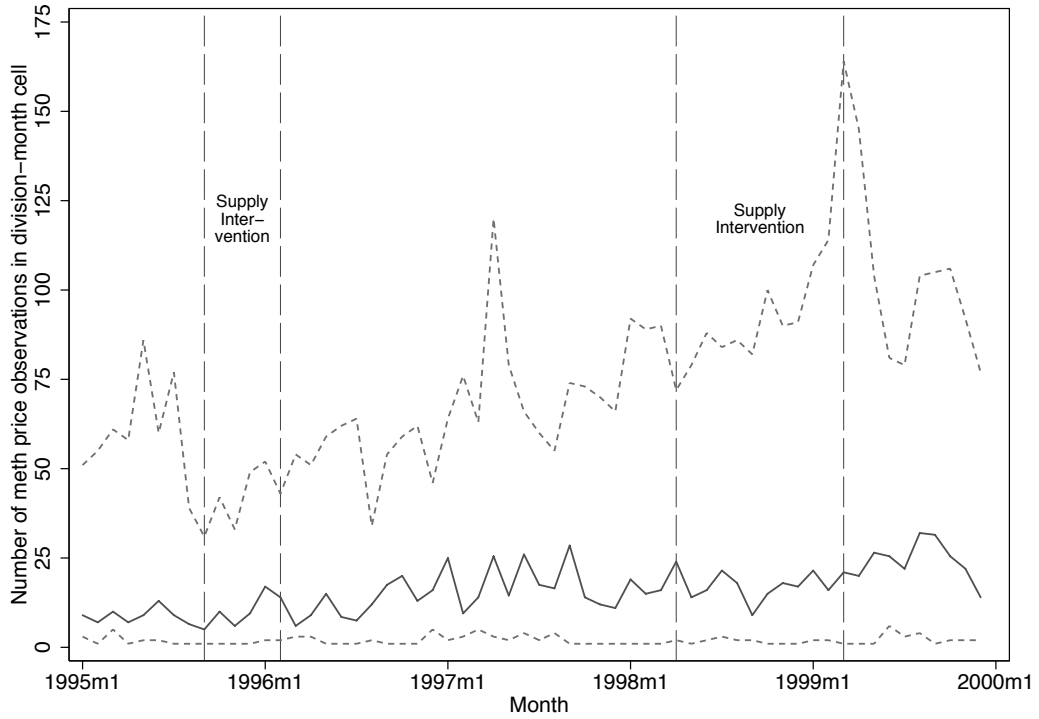

Notes: The lines show the minimum, median, and maximum number of meth price observations observed in states in a particular month.

Figure A2:   Construction of meth price instrumental variable as deviations of expected retail price of meth during interventions from overall trend lines, STRIDE, 1995–1999
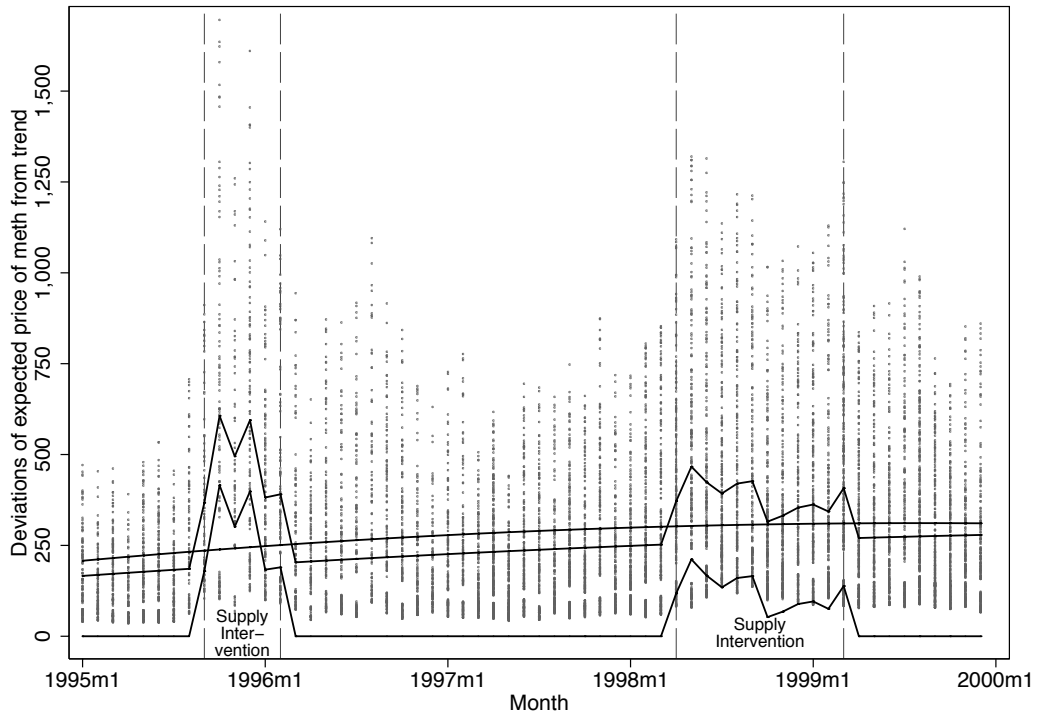
Notes: Authors' calculations from STRIDE. Dots represent individual observations for the expected price of pure meth. The smooth curve is the quadratic monthly time trend of expected meth prices. The bottom dark line is the instrumental variable—equal to zero outside of the supply interventions, and equal to the deviation off the trend during the intervention.

Figure A3:   Data quality analysis, TEDS and AFCARS, 1995–1999
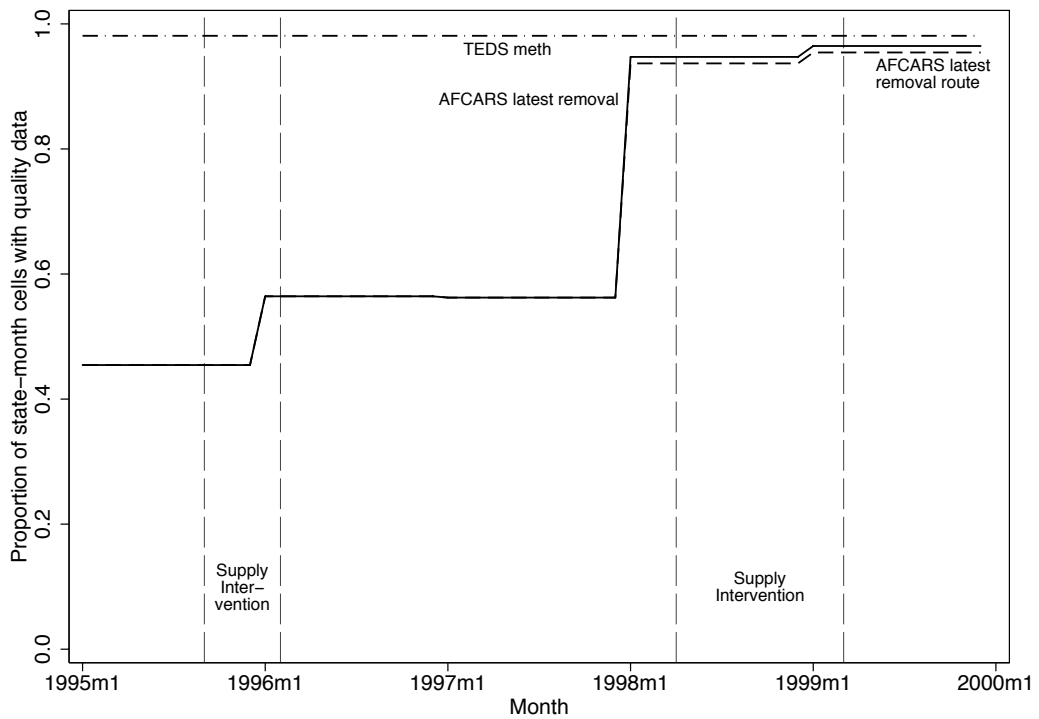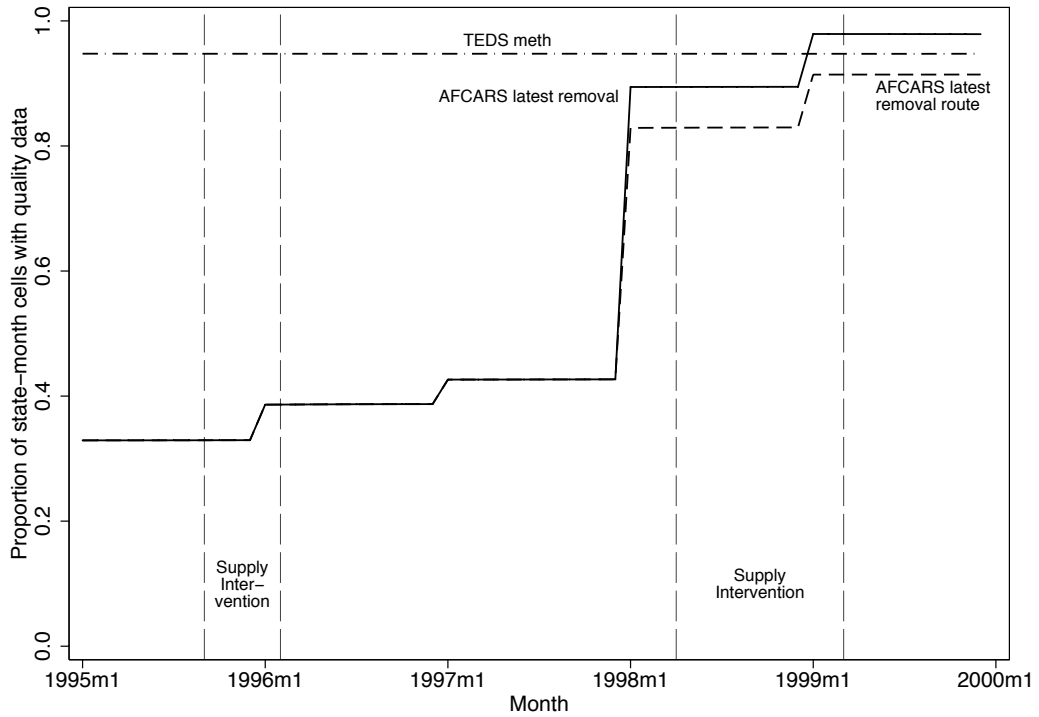
Notes: In the first figure, states are weighted by population aged 15–49. In the second figure, states are weighted by the number of TEDS patients in the last year of the sample who report meth use.