Evaluating the Impact of the Electronic Integration of Prescription Drug Monitoring Programs on Overdose Fatality Rates

Jonathan Magerko

Applied Econometrics II

The University of Akron

12/07/2025

Background

For this study, I look at the effect the electronic integration of Prescription Drug Monitoring Programs has had on the number of drug overdose fatalities per 100,000 people, at the state level.

Prescription Drug Monitoring Programs (PDMPs) are state-level databases that collect information on the prescribing and dispensing of controlled substances. The intention of these databases is to help identify potential misuse of prescriptions, to improve prescribing practices, and to ultimately reduce drug-related fatalities. The first PDMP was introduced in California in 1939, however, in 1991, Oklahoma became the first state to introduce the modern PDMP system that electronically collects and distributes prescription data. Most states have implemented this type of modern PDMP system within the last two decades. For a majority of these states, the modern PDMP system was their first time implementing PDMP legislation. (D'Souza, Lang, & Eldrige, 2024)

This electronic integration of PDMP systems is particularly important as it allows states to collect real-time data on information related to the movement of prescriptions, which can allow for a plethora of actions to be done, such as identifying cases of doctor shopping, forged prescriptions, and fraudulent prescribing practices. (Islam & McRae, 2014)

Understanding whether electronic integration reduces overdose deaths matters because as states continue to experience rising rates of prescription misuse and fatalities, policymakers need evidence about which interventions improve outcomes. If electronic PDMP integration significantly reduces fatal overdose rates, then policymakers have additional evidence which can be used to justify expanding funding for more robust systems. Conversely, if the electronic

integration of these systems doesn't significantly reduce fatal overdose rates, then it may suggest that other targeted interventions may be more effective.

Data and Methodology

This study uses state-level data including all states, except for Missouri (which formally implemented a PDMP program in late 2023, not enough data available to measure impact) which consists of information regarding drug overdose fatalities and PDMP electronic integration dates. The outcome variable is represented as the number of intentional and unintentional drug overdose deaths per 100,000 people in a given year and state. This data was collected from CDC WONDER (2000, 2003, 2017), which provides state-level mortality data using the ICD-9 codes for the 2000 and 2003 collections, and the ICD-10 codes for the 2017 collection. Each collection contains multiple years of data. Data from 1980-2023 was collected. However, the analysis uses data up until 2016 since its at this point that the last state in the study (Pennsylvania) electronically integrated their PDMP. The dates of electronic integration were collected from PDMP TTAC (2017), which reports a few dates relevant to PDMP's, including when each state began to receive prescription data electronically. Pennsylvania's date of electronic integration was collected from the Philadelphia County Medical Society (2016).

The outcome variable, as mentioned above, is the number of intentional and unintentional drug overdoses per 100,000 people in a given state and year.[1] Across 1,795 observations from 1980-2016, the mean overdose rate is 6.59 fatalities per 100,000 people, with a standard deviation of 5.84. This indicates substantial variation across states and time.[2]

[1] This rate (number of overdose fatalities / state population) was calculated by CDC Wonder.
[2] See table 1 in appendix for the full descriptive table.

The main regressor in this study is the difference-in-differences indicator variable. This variable is equal to 1 for a given state and year combination if the state has electronically integrated its PDMP; 0 otherwise. For states that implemented the integration between January and June, the DiD is set to 1 in that same year. For states that implemented the integration between July and December, the DiD is set to 1 for the following year. This approach accounts for the fact that late-year implementations wouldn't have had a full-year effect on the outcome.

Since states implemented electronic integration of PDMPs at different times, the treatment is considered staggered. For example, Oklahoma implemented electronic integration in 1991 whereas, for the purposes of this study, Pennsylvania was the last state to do so in 2016. About two-thirds of states have electronically integrated their PDMPs within the last two decades.[3]

In this study, I estimate four total two-way fixed effects difference-in-differences models (TWFE DiD). All models include state and year fixed effects and differ only by the form of the outcome variable (linear vs log-linear) and by the treatment of standard errors (non-clustered vs clustered at the state-level).

The linear specification is as follows:

$$OdRate_{st} = \beta_0 + \beta_1 PDMPIntegration_{st} + State_s + Year_t + \epsilon_{st} \qquad (1)$$

Where *OdRate* measures the number of drug fatalities per 100,000 people in state *s* and year *t. PDMPIntegration* is the difference-in-difference indicator equal to 1 if a state is considered to have electronically integrated their PDMP in a given

[3] See table 2 in appendix for the full table showing breakdown of implementation

year; 0 otherwise[4]. *State* and *Year* represent state and year fixed effects. This first equation is the basis for models 1 and 2. Model 2 has standard errors that are robust to heteroskedasticity and clustered at the state-level whereas model 1 only has standard errors that are robust to heteroskedasticity.

The log-linear specification is as follows:

$$Log(OdRate_{st}) = \beta_0 + \beta_1 PDMPIntegration_{st} + State_s + Year_t + \epsilon_{st} \quad (2)$$

Where *Log(OdRate)* is the natural-log of the number of drug fatalities per 100,000 people. Other variables remain the same as in the first equation. This second equation forms the basis for models 3 and 4. The only difference between models 3 and 4 is the treatment of the robust standard errors, similar to that of models 1 and 2.

Results and Discussion

According to model 2, holding everything else constant and on average, the introduction of an electronically integrated PDMP is associated with a 0.0083-unit increase in drug overdose fatalities per 100,000 people, compared to what it would have been if there hadn't been a treatment.

According to model 4, holding everything else constant and on average, the introduction of an electronically integrated PDMP is associated with an estimated 0.35% increase in drug overdose fatalities per 100,000 people, similarly, compared to what it would have been if there hadn't been a treatment.

[4] See paragraph 3 in the data and methodology section to see adjusted DiD indicator creation

In terms of statistical significance, the treatment is not considered to be significant at any conventional level in either model. In both model specifications, the clustering of robust standard errors increases the size of the standard errors, but the coefficients were already insignificant.

Given the lack of statistical significance, the estimated effect of the treatment does not provide credible evidence of an economically meaningful impact.

While the coefficients for the treatment are near-zero and insignificantly positive (PDMP introduction is expected to reduce prescription misuse), this could be due in part to omitted variable bias. While these models account for state (time-invariant differences between states) and year (time-variant effects that impact each state) fixed effects, they do not account for time-variant, state-specific factors. For instance, as the opioid crisis has unfolded, each state may have implemented other pieces of legislation or different types of interventions at different times. Another factor is that I included both intentional and unintentional drug overdose fatalities under the belief that the effectiveness of PDMPs would impact the accessibility of prescriptions for all individuals at risk, regardless of intent. Analyzing these outcomes separately may provide a clearer understanding of how the effectiveness of PDMPs affect different types of overdoses.

Overall, based on the results of the TWFE DiD models, there is no statistically or economically significant evidence that the electronic integration of Prescription Drug Monitoring Programs reduces overdose fatalities at the state-level.

Appendix: Tables

Table 1: Summary Statistics of the Outcome Variable

Variable	N	Mean	Std. Dev.	Min	Max
Overdose Fatalities (per 100,000 people)	1795	6.5929	5.8378	0.02	46.58

Sources: Own calculations.
Notes: N represents the number of state-year observations used in the analysis. The outcome variable measures the total unintentional and intentional overdose fatalities per 100,000 people in a given state and year.

Table 2: Breakdown of Treatment Implementation

Year	States that electronically integrated their PDMP
1991	OK
1992	HI
1994	MA
1996	UT
1997	NV
1998	IN
1999	KY, NY
2000	IL
2001	TX
2002	WV
2003	MI, VA
2004	ID, ME, WY
2005	NM
2006	AL, RI, OH, TN
2007	CA, CO, NC, ND
2008	MS, SC, CT, AZ, LA
2009	IA, VT
2010	MN
2011	KS, NE, OR, AK, FL, NJ, WA, SD
2012	DE, MT
2013	AR, WI, GA, MD
2014	NH
2016	PA

Sources: PDMP TTAC (2017), PCMS (2017), and own calculations.
Notes: Missouri is excluded. This table is intended to give an overview of the staggered adoption pattern.

Table 3: Impact of Electronic Integration of PDMP's on Drug Overdose Fatalities

Regressors	Model 1	Model 2	Model 3	Model 4
PDMP Electronic Integration	0.0083	0.0083	0.0035	0.0035
	(0.2552)	(0.5881)	(0.0413)	(0.0809)
Intercept	18.7937***	18.7937***	2.7810***	2.7810***
	(1.1290)	(1.2299)	(0.1022)	(0.1011)
Number of Obs.	1,795	1,795	1,795	1,795
Adjusted R-Square	0.7840	0.7840	0.7141	0.7141
Overall Significance	74.98***	40.44***	133.31***	82.57***
State and Year Fixed Effects?	Yes	Yes	Yes	Yes
State-Level Standard Errors?	No	Yes	No	Yes
Outcome Type	Linear	Linear	Log	Log

Sources: CDC WONDER (2000, 2003, 2017), PDMP TTAC (2017), PCMS (2016), and own calculations

Notes: Robust standard errors are in parentheses. Models 1 and 3 have non-clustered robust standard errors whereas models 2 and 4 have state-level clustered robust standard errors. *, **, and *** indicate 10%, 5%, and 1% significance levels, respectively. Models 1 and 2 use the linear form of the outcome variable whereas models 3 and 4 use the natural log form. The outcome variable is measured in units of drug overdose fatalities per 100,000 people in a given state and year.

Appendix: References

D'Souza, R. S., Lang, M., & Eldrige, J.S. (2024) *Prescription Drug Monitoring Program*. In StatPearls. StatPearls Publishing. Available at: https://www.ncbi.nlm.nih.gov/books/NBK532299/.

Islam, M. M., & McRae, I. S. (2014). *An inevitable wave of prescription drug monitoring programs in the context of prescription opioids: pros, cons and tensions.* BMC pharmacology & toxicology, 15, 46. Available at: https://doi.org/10.1186/2050-6511-15-46

PDMP Training and Technical Assistance Center (2017) *State PDMP implementation dates*. Brandeis University. Available at: https://pdaps.org/datasets/pdmp-implementation-dates, data extracted on October 25, 2025.

Philadelphia County Medical Society. (2016). *The PA PDMP system officially launches August 25!* Available at: https://philamedsoc.org/the-pa-pdmp-system-officially-launches-august-25/

Centers for Disease Control and Prevention, National Center for Health Statistics. (2000). *National Vital Statistics System, Mortality: Compressed Mortality File 1979-1998*. CDC WONDER Online Database. Compiled from Compressed Mortality File 1968-1988, Series 20, No. 2A. Available at: http://wonder.cdc.gov/cmf-icd9.html, data extracted on October 25, 2025.

Centers for Disease Control and Prevention, National Center for Health Statistics. (2003). *National Vital Statistics System, Mortality: Compressed Mortality File 1979-1998*. CDC WONDER Online Database. Compiled from Compressed Mortality File 1989-1998, Series 20, No. 2E. Available at: http://wonder.cdc.gov/cmf-icd9.html, data extracted on October 25, 2025.

Centers for Disease Control and Prevention, National Center for Health Statistics (2017). *National Vital Statistics System, Mortality: Compressed Mortality File 1999-2016*. CDC WONDER Online Database. Compiled from Compressed Mortality File 199-2016, Series 20, No. 2U. Available at http://wonder.cdc.gov/cmf-icd10.html, data extracted on October 25, 2025.

Appendix: SAS Codes

```
/* Jonathan Magerko

      Project 3 Code



/*




/* Data Importing */

proc import datafile =
"/home/u64112853/MySAS/JonathanMagerko_ProposedPanel_Project3.xlsx"

      out = work.db1_original

      dbms = xlsx

      replace;

      sheet = "Panel Data";

      getnames = y;

run;



proc import datafile =
"/home/u64112853/MySAS/JonathanMagerko_ProposedPanel_Project3.xlsx"

      out = work.db2_original

      dbms = xlsx

      replace;
```

```
      sheet = "Electronic Integrations";
      getnames = y;
run;

/* Data Cleaning */

/* rename year*/
data db2;
    set db2_original (rename=("Date of Electronic
Integration"n=FirstTreatmentYear));
    treatment_year = year(FirstTreatmentYear);
run;

data db1;
    set db1_original;
    /*removing later years (specified in proposal)*/
    if not (2017 le Year le 2023);
      /* removing suppressed observations due to confidentiality (18 total)
            then converting to numeric
                  */
      if outcome = "Suppressed" then outcome_clean = .;
            else outcome_clean = input(outcome, 8.2);
```

```
      drop outcome;
      rename outcome_clean = Outcome;
run;



data Combined;
      merge db1 db2;
      by state;
      /* Extract the year and month from the treatment date */
    treatment_year = year(FirstTreatmentYear);
    treatment_month = month(FirstTreatmentYear);
    /* Adjust the treatment year based on the month the treatment was
administered
            this is due to some treatments being administered late in the
year  */
      if treatment_month le 6 then adj_treatment_year = treatment_year;
            else adj_treatment_year = treatment_year + 1;
      /* DiD Indicator Creation */
      if year ge adj_treatment_year then DID=1;
            else DID=0;
      /* Creating log-form of outcome variable */
      log_outcome = log(outcome);
```

```
      keep state year outcome log_outcome DID treatment_year;
run;

proc sort data = combined;
      by state;
run;

/* Values for Summary Statistics */

proc means data = combined;
      var outcome;
run;

proc freq data = combined;
      table DID;
run;

/* DiD Model Setup */

/* linear, non-clustered standard errors */
```

```
ods output ParameterEstimates=PEforModel1 DataSummary=ObsModel1
                  FitStatistics=AdjRsqModel1 Effects=OverallSigModel1;
proc surveyreg data=combined plots=none;
      class state year;
      model1: model outcome=DID state year/solution adjrsq;
run;

/* linear, state-level clustered standard errors */
ods output ParameterEstimates=PEforModel2 DataSummary=ObsModel2
                  FitStatistics=AdjRsqModel2 Effects=OverallSigModel2;
proc surveyreg data=combined plots=none;
      class state year;
      cluster state;
      model2: model outcome=DID state year/solution adjrsq;
run;

/* log-form, non-clustered standard errors */
ods output ParameterEstimates=PEforModel3 DataSummary=ObsModel3
                  FitStatistics=AdjRsqModel3 Effects=OverallSigModel3;
proc surveyreg data=combined plots=none;
      class state year;
```

```
      model3: model log_outcome=DID state year/solution adjrsq;
run;

/* log-form, clustered standard errors */
ods output ParameterEstimates=PEforModel4 DataSummary=ObsModel4
                  FitStatistics=AdjRsqModel4 Effects=OverallSigModel4;
proc surveyreg data=combined plots=none;
      class state year;
      cluster state;
      model4: model log_outcome=DID state year/solution adjrsq;
run;

/* Final Results Table */
data table_long;
      length Model $10; /* Makes sure the variable Model has the right length
and its values are not truncated */
      length Parameter $30; /* Makes sure the variable Parameter has the
right length and its values are not truncated */
      set PEforModel1  PEforModel2 PEforModel3 PEforModel4 indsname=M;
/*"indsname" creates an indicator variable (here I call it "M") that tracks
the name of databases use in the "set" statement */
      *ThisIsM=M;
```

```
      where (Estimate ne 0) and (substr(Parameter,1,5) ne "State") and
(substr(Parameter,1,4) ne "Year");


      if    M="WORK.PEFORMODEL1" then Model="Model1";
            else if M="WORK.PEFORMODEL2" then Model="Model2";
            else if M="WORK.PEFORMODEL3" then Model="Model3";
            else Model = "Model 4";


      length star $3;
      if probt le 0.01 then Star="***";
            else if probt le 0.05 then star="**";
            else if probt le 0.1 then star="*";
            else star="";



      Results=Estimate;
      EditedResults=cats(put(Results,comma16.4),star);
      Output;


      Results=StdErr;
      EditedResults=cats("(",put(Results,comma16.4),")");
      Output;
```

```
      keep Model Parameter EditedResults;
run;

proc sort data=Table_Long out=Table_Long_Sorted;
      by Model Parameter;
run;

/* Creating Results Columns */

data Model1Results (rename=(EditedResults=Model1))
       Model2Results (rename=(EditedResults=Model2))
       Model3Results (rename=(EditedResults=Model3))
       Model4Results (rename=(EditedResults=Model4));
      set table_long_sorted;
      if Model="Model1" then output Model1Results;
            else if Model="Model2" then output Model2Results;
            else if Model="Model3" then output Model3Results;
            else output Model4Results;
      drop Model;
run;
```

```
proc sort data = Model1Results;
      by Parameter;
run;

proc sort data = Model2Results;
      by Parameter;
run;

proc sort data = Model3Results;
      by Parameter;
run;

proc sort data = Model4Results;
      by Parameter;
run;

/* Final Results Table Steps*/

data table_wide;
      merge Model1Results Model2Results Model3Results Model4Results;
```

```
      by Parameter;

      Length Order 3;

      if Parameter="DID" then Order =1;
            else if Parameter="Intercept " then Order =2;
            else order =3;

      if mod(_n_,2)=1 then Regressors=Parameter;

run;

proc sort data=table_wide out=table_wide_sorted(drop=Order Parameter);
      by Order;
run;

/* Add other statistics to the table */

/* Num of Obs. */
data NumofObs;
      merge ObsModel1 (rename=(NValue1=NVModel1) drop=CValue1)
```

```
                ObsModel2 (rename=(NValue1=NVModel2) drop=CValue1)
                ObsModel3 (rename=(NValue1=NVModel3) drop=CValue1)
                ObsModel4 (rename=(NValue1=NVModel4) drop=CValue1);
        where Label1="Number of Observations";
        Model1=put(NVModel1,comma16.);
        Model2=put(NVModel2,comma16.);
        Model3=put(NVModel3,comma16.);
        Model4=put(NVModel4,comma16.);
        keep Label1 Model1 Model2 Model3 Model4;
run;

/* Adjusted Rsq. */
data AdjRsq;
        merge AdjRsqModel1 (rename=(Cvalue1=Model1))
                AdjRsqModel2 (rename=(Cvalue1=Model2))
                AdjRsqModel3 (rename=(Cvalue1=Model3))
                AdjRsqModel4 (rename=(Cvalue1=Model4));
        where Label1="Adjusted R-Square";
        keep Label1 Model1 Model2 Model3 Model4;
run;
```

```
/* F-Test for Overall Sig. */
data OSM1 (rename=(EditedValue=Model1))
       OSM2 (rename=(EditedValue=Model2))
       OSM3 (rename=(EditedValue=Model3))
       OSM4 (rename=(EditedValue=Model4));
      set OverallSigModel1 OverallSigModel2
            OverallSigModel3 OverallSigModel4
            indsname=M;
      *ThisIsM=M;
      where Effect="Model";
      Label1="Overall Significance";

      Length Star $3;
      if ProbF le 0.01 then Star="***";
            else if ProbF le 0.05 then Star="**";
            else if ProbF le 0.1 then Star="*";
            else Star="";

      EditedValue=Cats(Put(Fvalue,comma20.2),star);

      if M="WORK.OVERALLSIGMODEL1" then output OSM1;
```

```
            else if M="WORK.OVERALLSIGMODEL2" then output OSM2;
            else if M="WORK.OVERALLSIGMODEL3" then output OSM3;
            else output OSM4;



      keep Label1 EditedValue;
run;

data OverallSig;
      merge OSM1 OSM2 OSM3 OSM4;
run;

/* Combining these "other statistics" */
data OtherStat;
      Set NumofObs AdjRsq OverallSig;
      rename Label1=Regressors;
run;

/* Descriptive Indicators */
data OtherInfo;
      length  Regressors Model1 Model2 Model3 Model4 $50;
```

```
Regressors="State and Year Fixed Effects?";

Model1="Yes";

Model2="Yes";

Model3="Yes";

Model4="Yes";

output;



Regressors="State-Level Standard Errors?";

Model1="No";

Model2="Yes";

Model3="No";

Model4="Yes";

output;



Regressors="Outcome Type";

Model1="Linear";

Model2="Linear";

Model3="Log";

Model4="Log";

output;
```

```
run;


/* Adding these "other statistics" to the final results table */
data table_wide_sorted_withStats;
      Set table_wide_sorted OtherStat OtherInfo;
run;


/* Labeling */
proc format;
      value $VariableName(default=50)
            "DID" = "PDMP Electronic Integration"
            "Number of Observations" = "Number of Obs."
      ;
run;


/* Outputting Final Results Table */
ods excel
file="/home/u64112853/MySAS/JonathanMagerko_Project3_ResultsTable.xlsx"
options(Embedded_Titles="ON" Embedded_Footnotes="ON");
Title "Table 3: Impact of Electronic Integration of PDMP's on Drug Overdose
Fatalities";
```

```
Footnote1 J=L "Sources: CDC WONDER (2000, 2003, 2017), PDMP TTAC (2017), PCMS
(2016), and own calculations";

Footnote2 J=L "Notes: Robust standard errors are in parentheses. Models 1 and
3 have non-clustered robust standard

                        errors whereas models 2 and 4 have state-level
clustered robust standard errors.

                        *, **, and *** indicate 10%, 5%, and 1% significance
levels, respectively.

                        Models 1 and 2 use the linear form of the outcome
variable whereas models 3 and 4 use the natural log form.

                        The outcome variable is measured in units of drug
overdose fatalities per 100,000 people in a given state and year.";

proc print data=Table_Wide_Sorted_withStats noobs;

       Var Regressors;

       Var Model1 Model2 Model3 Model4 / style(header)={Just=Center}
style(data)={Just=Center tagattr="type:String"};

       format Regressors $VariableName.;

run;

ods excel close;
```