

# Protecting Confidentiality in Public Data Reporting:

## *An Application of a Statistical Jitter Function*

*June 2015*

**Daniel Webster, MSW, PhD**

**Bill Dawson, MSW**

**Michael Armijo**

California Child Welfare Indicators Project (CCWIP)

School of Social Welfare

University of California, Berkeley

*CCWIP is a collaboration of the California Department of Social Services and the School of Social Welfare, University of California at Berkeley, and is supported by the California Department of Social Services and the Stuart Foundation*

*We would like to acknowledge the significant contribution of Partners for Our Children, School of Social Work, University of Washington in our development of this approach*



STUART FOUNDATION



Berkeley Social Welfare  
UC BERKELEY SCHOOL OF SOCIAL WELFARE



# DHCS Public Reporting

- Per “Public Aggregate Reporting – DHCS Business Reports Guidelines”:
- In order to achieve the goals of public reporting and protection of personally identifiable information, “procedures that appropriately and accurately de-identify data when publicly reporting are necessary.”

# Small cells

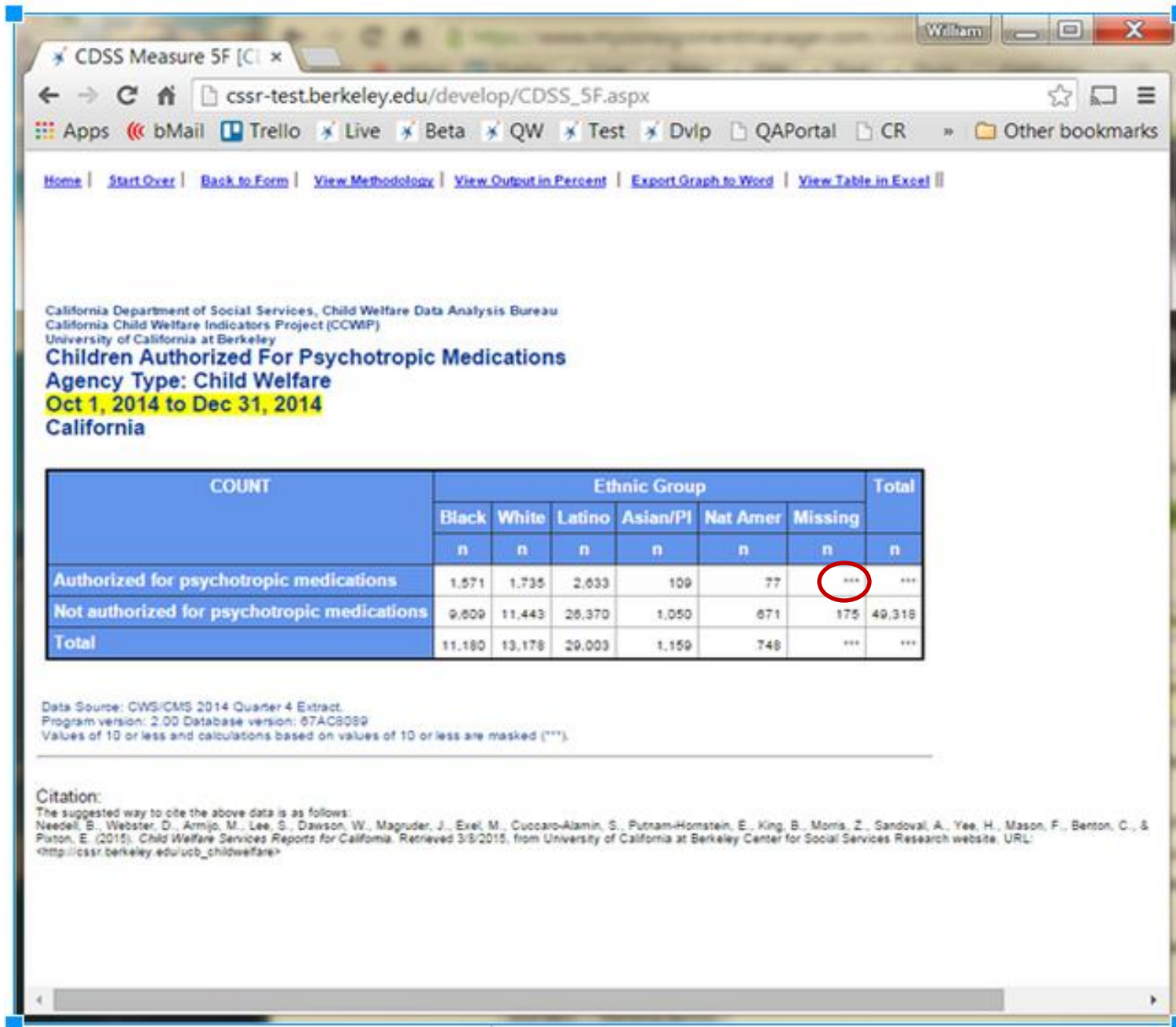
- Per *Public Aggregate Reporting – DHCS Business Reports Guidelines* (pg. 11 of 41):
- “Cell suppression includes two parts: suppression of the cell that has a **value less than 11** and suppression of other **complementary cells that may allow for calculation** of the cell being suppressed. Complementary cells may have larger cell sizes that support calculation of the exact number of cases in cells of less than 11 by subtraction or other mathematical means.”\*

\*Emphasis added

# DHCS guidelines applied to 5F

- Applying these DHCS guidelines to the public report, *5F Children Authorized for Psychotropic Medications*, using an approach of masking (\*\*\*) the small and complementary cells suppresses a lot of data.

# An early masking attempt (no longer online) ...



CDSS Measure 5F [C]

cssr-test.berkeley.edu/develop/CDSS\_5F.aspx

Apps bMail Trello Live Beta QW Test Dvlp QAPortal CR » Other bookmarks

[Home](#) | [Start Over](#) | [Back to Form](#) | [View Methodology](#) | [View Output in Percent](#) | [Export Graph to Word](#) | [View Table in Excel](#)

California Department of Social Services, Child Welfare Data Analysis Bureau  
California Child Welfare Indicators Project (CCWIP)  
University of California at Berkeley

**Children Authorized For Psychotropic Medications**  
**Agency Type: Child Welfare**  
**Oct 1, 2014 to Dec 31, 2014**  
**California**

COUNT	Ethnic Group						Total
	Black	White	Latino	Asian/PI	Nat Amer	Missing	
	n	n	n	n	n	n	n
Authorized for psychotropic medications	1,571	1,735	2,633	109	77	***	***
Not authorized for psychotropic medications	9,609	11,443	26,370	1,050	671	175	49,318
Total	11,180	13,178	29,003	1,159	748	***	***

Data Source: CWS/CMS 2014 Quarter 4 Extract.  
Program version: 2.00 Database version: 67AC8089  
Values of 10 or less and calculations based on values of 10 or less are masked (\*\*\*).

Citation:  
The suggested way to cite the above data is as follows:  
Needell, B., Webster, D., Armijo, M., Lee, S., Dawson, W., Magruder, J., Exell, M., Cuccaro-Alamin, S., Putnam-Hornstein, E., King, B., Morris, Z., Sandoval, A., Yee, H., Mason, F., Benton, C., & Putnam, E. (2015). *Child Welfare Services Reports for California*. Retrieved 3/8/2015, from University of California at Berkeley Center for Social Services Research website, URL: [http://cssr.berkeley.edu/ucb\\_childwelfare/](http://cssr.berkeley.edu/ucb_childwelfare/)

[Home](#) | [Start Over](#) | [Back to Form](#) | [View Methodology](#) | [View Table in Excel](#) ||

California Department of Social Services, Child Welfare Data Analysis Bureau  
California Child Welfare Indicators Project (CCWIP)  
University of California at Berkeley

## Children Authorized For Psychotropic Medications

Agency Type: Child Welfare

Oct 1, 2014 to Dec 31, 2014

County	In care	Authorized for psychotropic medication	%
California	55,452	6,134	11.1%
Alameda	1,193	84	7.0%
Alpine	***	***	***
Amador	65	12	18.5%
Butte	533	63	11.8%
Calaveras	147	***	***
Colusa	26	***	***
Contra Costa	1,014	67	6.6%
Del Norte	103	***	***



CDSS Measure 5F [C]

cssr-test.berkeley.edu/develop/CDSS\_5F.aspx

Apps

bMail

Trello

Live

Beta

QW

Test

Dvlp

QAPortal

CR

Other bookmarks

Home

Start Over

Back to Form

View Methodology

View Output in Percent

Export Graph to Word

View Table in Excel

California Department of Social Services, Child Welfare Data Analysis Bureau

California Child Welfare Indicators Project (CCWIP)

University of California at Berkeley

Children Authorized For Psychotropic Medications

Agency Type: Child Welfare

Oct 1, 2014 to Dec 31, 2014

California

COUNT	Ethnic Group						Total
	Black	White	Latino	Asian/PI	Nat Amer	Missing	
	n	n	n	n	n	n	n
Authorized for psychotropic medications	1,571	1,735	2,633	109	77	9	6,134
Not authorized for psychotropic medications	9,609	11,443	26,370	1,050	671	175	49,318
Total	11,180	13,178	29,003	1,159	748	184	55,452

Data Source: CWS/CMS 2014 Quarter 4 Extract.

Program version: 2.00 Database version: 67AC8089

Values of 10 or less and calculations based on values of 10 or less are masked (\*\*\*).

Citation:

The suggested way to cite the above data is as follows:

Needell, B., Webster, D., Armijo, M., Lee, S., Dawson, W., Magruder, J., Exel, M., Cuccaro-Alamin, S., Putnam-Hornstein, E., King, B., Morris, Z., Sandoval, A., Yee, H., Mason, F., Benton, C., & Putnam, E. (2015). Child Welfare Services Reports for California. Retrieved 3/8/2015, from University of California at Berkeley Center for Social Services Research website, URL: <http://cssr.berkeley.edu/ucb\_childwelfare>

## Masking Solution Currently Online ...

California Department of Social Services, Child Welfare Data Analysis Bureau

California Child Welfare Indicators Project (CCWIP)

University of California at Berkeley

### Children Authorized For Psychotropic Medications

Agency Type: Child Welfare

Oct 1, 2014 to Dec 31, 2014

California

COUNT	Ethnic Group						Total
	Black	White	Latino	Asian/PI	Nat Amer	Missing	n
	n	n	n	n	n	n	
Authorized for psychotropic medications	1,571	1,735	2,633	109	***	***	6,134
Not authorized for psychotropic medications	9,609	11,443	26,370	1,050	671	175	49,318
Total	11,180	13,178	29,003	1,159	***	***	55,452

Data Source: CWS/CMS 2014 Quarter 4 Extract

Program version: 2.00 Database version: 67AC8089

Values of 10 or less and calculations based on values of 10 or less are masked (\*\*\*).

Additional values (the lowest available) are masked to prevent calculation of values of 10 or less.



## Children Authorized For Psychotropic Medications

Agency Type: Child Welfare

Oct 1, 2014 to Dec 31, 2014

County	In care	Authorized for psychotropic medication	%
California	55,452	6,134	11.1%
Alameda	1,193	84	7.0%
Alpine	xxx	xxx	xxx
Amador	65	12	18.5%
Butte	533	63	11.8%
Calaveras	147	xxx	xxx
Colusa	26	xxx	xxx
Contra Costa	1,014	67	6.6%
Del Norte	103	xxx	xxx
El Dorado	296	xxx	xxx
Fresno	1,959	243	12.4%
Glenn	93	12	12.9%
Humboldt	316	28	8.9%
Imperial	426	57	13.4%
Inyo	14	xxx	xxx
Kern	1,375	143	10.4%
Kings	483	30	6.2%
Lake	110	18	16.4%
Lassen	57	xxx	xxx
Los Angeles	18,412	1,908	10.4%
Madera	327	17	5.2%
Marin	88	xxx	xxx
Mariposa	11	xxx	xxx
Mendocino	245	29	11.8%

# **Children Authorized For Psychotropic Medications**

**Agency Type: Child Welfare**

**Oct 1, 2014 to Dec 31, 2014**

**Contra Costa**

COUNT	Placement Type														Total
	Pre-Adopt	Kin	Foster	FFA	Court Specified	Group	Shelter	Non-FC	Transitional Housing	Guardian-Dependent	Runaway	Trial Home Visit	Other	Missing	
	n	n	n	n	n	n	n	n	n	n	n	n	n	n	
Authorized for psychotropic medications	xxx	xxx	xxx	20	xxx	35	xxx	xxx	xxx	xxx	xxx	xxx	xxx	xxx	67
Not authorized for psychotropic medications	47	354	170	240	xxx	68	xxx	xxx	xxx	39	xxx	xxx	18	xxx	947
Total	xxx	xxx	xxx	260	xxx	103	xxx	xxx	xxx	xxx	xxx	xxx	xxx	xxx	1,014

[http://cssr.berkeley.edu/ucb\\_childwelfare/CDSS\\_5F.aspx](http://cssr.berkeley.edu/ucb_childwelfare/CDSS_5F.aspx)

# The “Jitter” Function\*

- Goals:
  - Prevent the identification of individuals
  - Preserve general trends and the overall integrity of the data

\* We would like to acknowledge the significant contribution of Partners for Our Children, School of Social Work, University of Washington in our development of this approach.

# The “Jitter” Function

- A random number is added to each non-zero cell in each table that is returned. The random numbers are generated from a specific probability distribution.
- Zero values remain as zero. This preserves the presence/absence of an event.
- Jittered results are set to 1 if the jitter function reduces them to a value less than 1.

\* We would like to acknowledge the significant contribution of Partners for Our Children, School of Social Work, University of Washington in our development of this approach.

# The Jitter Function: Potential Concerns

- Output tables, rather than underlying data are jittered.
- As a result, output tables in which totals would match with non-jittered data, may differ slightly

COUNT	Gender			Total
	Female	Male	Missing	
	n	n	n	
Authorized for psychotropic medications	31	38	0	69
Not authorized for psychotropic medications	502	445	0	947
Total	533	483	0	1,016

COUNT	Age Group							Total
	Under 1	1-2	3-5	6-10	11-15	16-17	Missing	
	n	n	n	n	n	n	n	
Authorized for psychotropic medications	0	0	0	2	33	35	0	70
Not authorized for psychotropic medications	63	146	149	264	225	104	0	951
Total	63	146	149	266	258	139	0	1,021

COUNT	Ethnic Group						Total
	Black	White	Latino	Asian/PI	Nat Amer	Missing	
	n	n	n	n	n	n	
Authorized for psychotropic medications	39	14	12	0	0	0	65
Not authorized for psychotropic medications	397	290	218	29	12	0	950
Total	436	304	230	29	12	0	1,015

COUNT	Placement Type														Total
	Pre-Adopt	Kin	Foster	FFA	Court Specified	Group	Shelter	Non-FC	Transitional Housing	Guardian-Dependent	Runaway	Trial Home Visit	Other	Missing	
	n	n	n	n	n	n	n	n	n	n	n	n	n	n	
Authorized for psychotropic medications	0	5	2	21	0	34	0	0	1	1	1	0	1	0	66
Not authorized for psychotropic medications	46	354	170	240	0	68	0	0	1	39	10	0	19	0	947
Total	46	359	172	261	0	102	0	0	2	40	11	0	20	0	1,013

- With jittered data, county totals differ slightly from table to table

# The Jitter Function: Potential Concerns

- Output tables, rather than underlying data are jittered.
- As a result, output tables in which totals would match with non-jittered data, may differ slightly
  - **This issue can be addressed in presentation and documentation of the jitter function**

[Non-public site with jittered data](#)



# The Jitter Function: Potential Concerns

- A malicious user could attempt to determine the true value of certain cells by averaging multiple runs of the same query.
  - However, for a given extract, our system returns the same (jittered) results each time the same query is called.
  - Averaging values returned from different views of the same data won't necessarily return the true value. Even if it does, uncertainty remains.

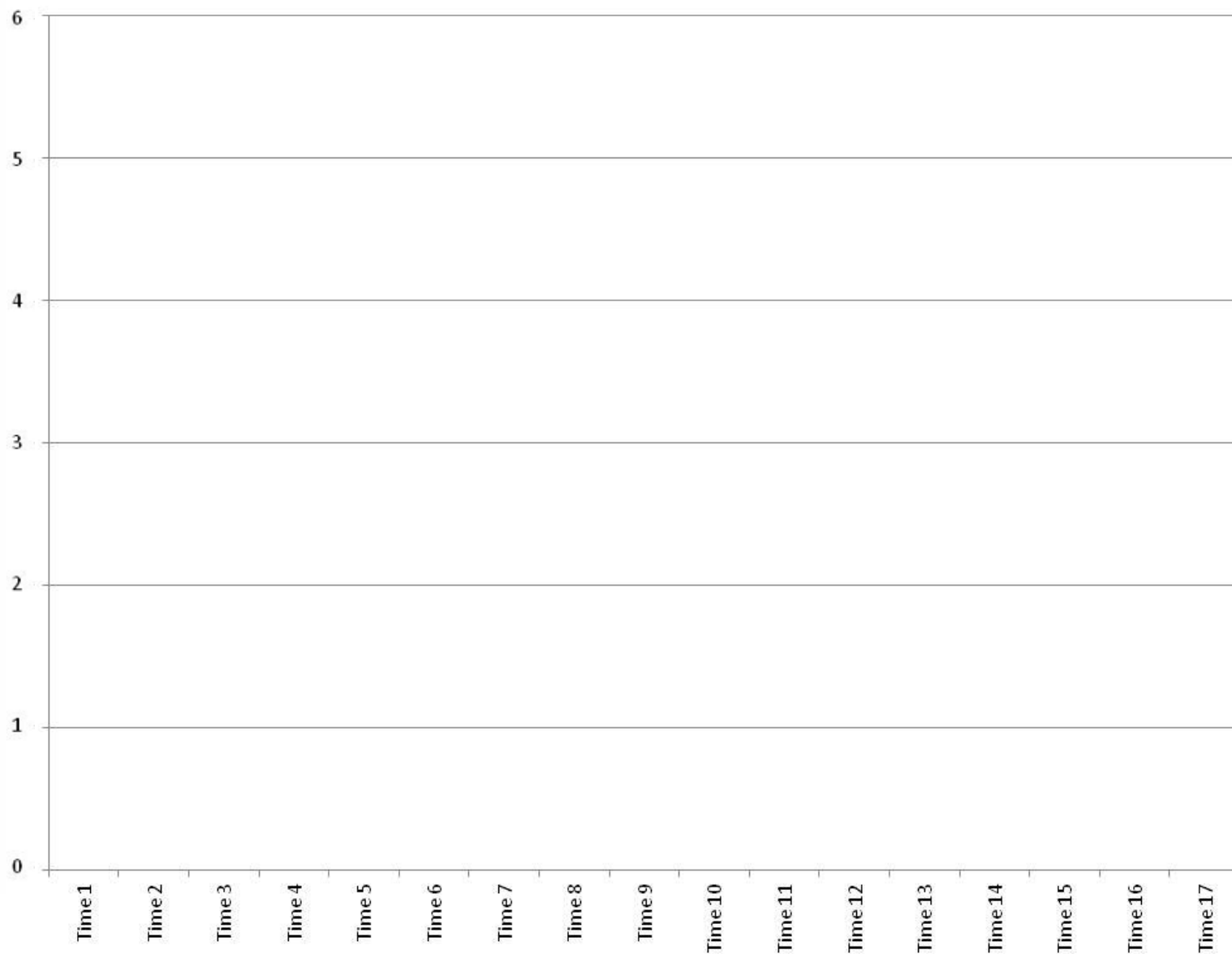
# The Jitter Function: Potential Concerns

- To provide additional anonymity, the specifics of the jitter function can be kept confidential: the type of the random distribution, standard deviation, and other details.

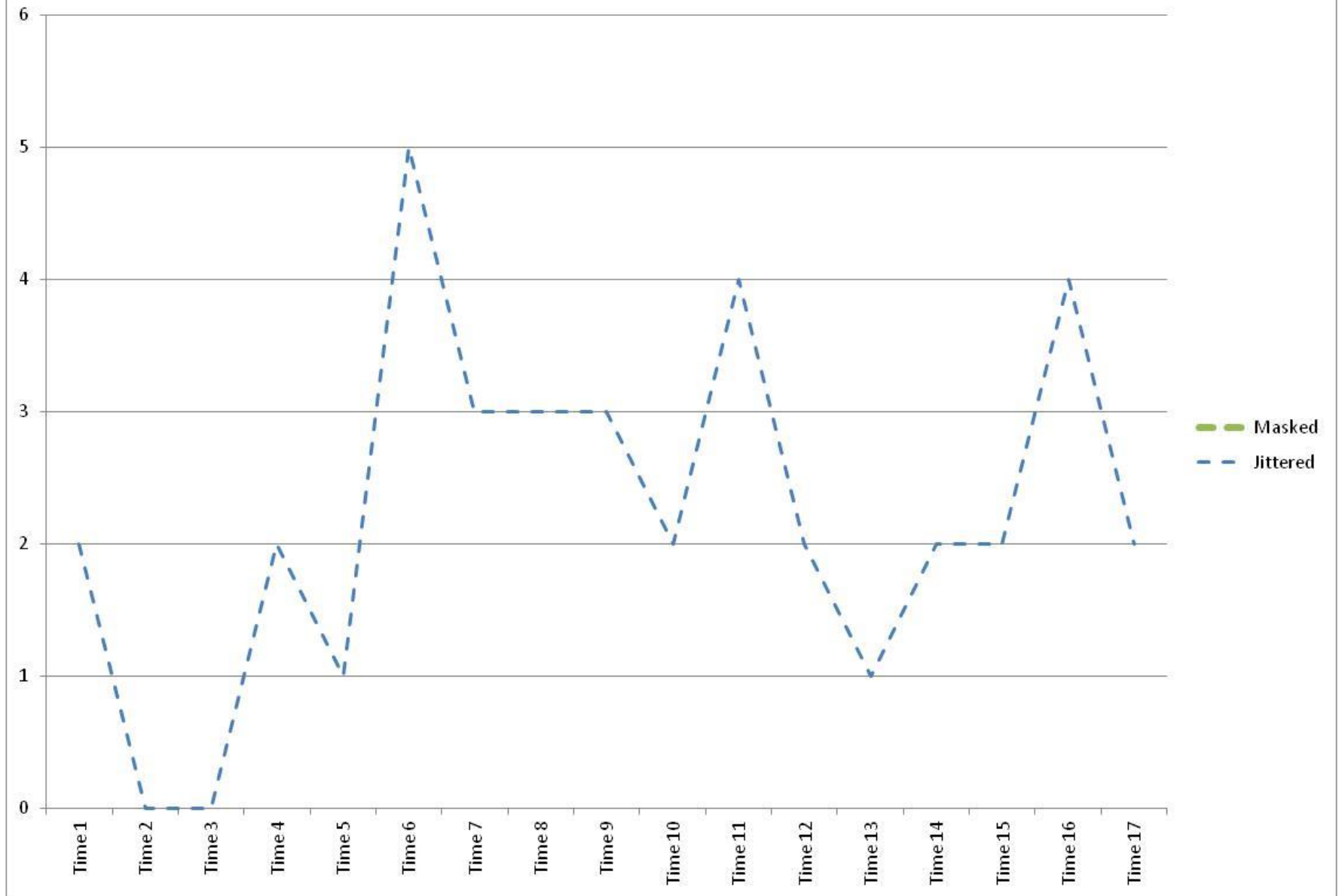
# Sample Effects of Jitter Function

- The following slides show available numerators in three versions of the data:
  - (1) With values masked (as currently on the live site)
  - (2) Jittered values
  - (3) Actual values

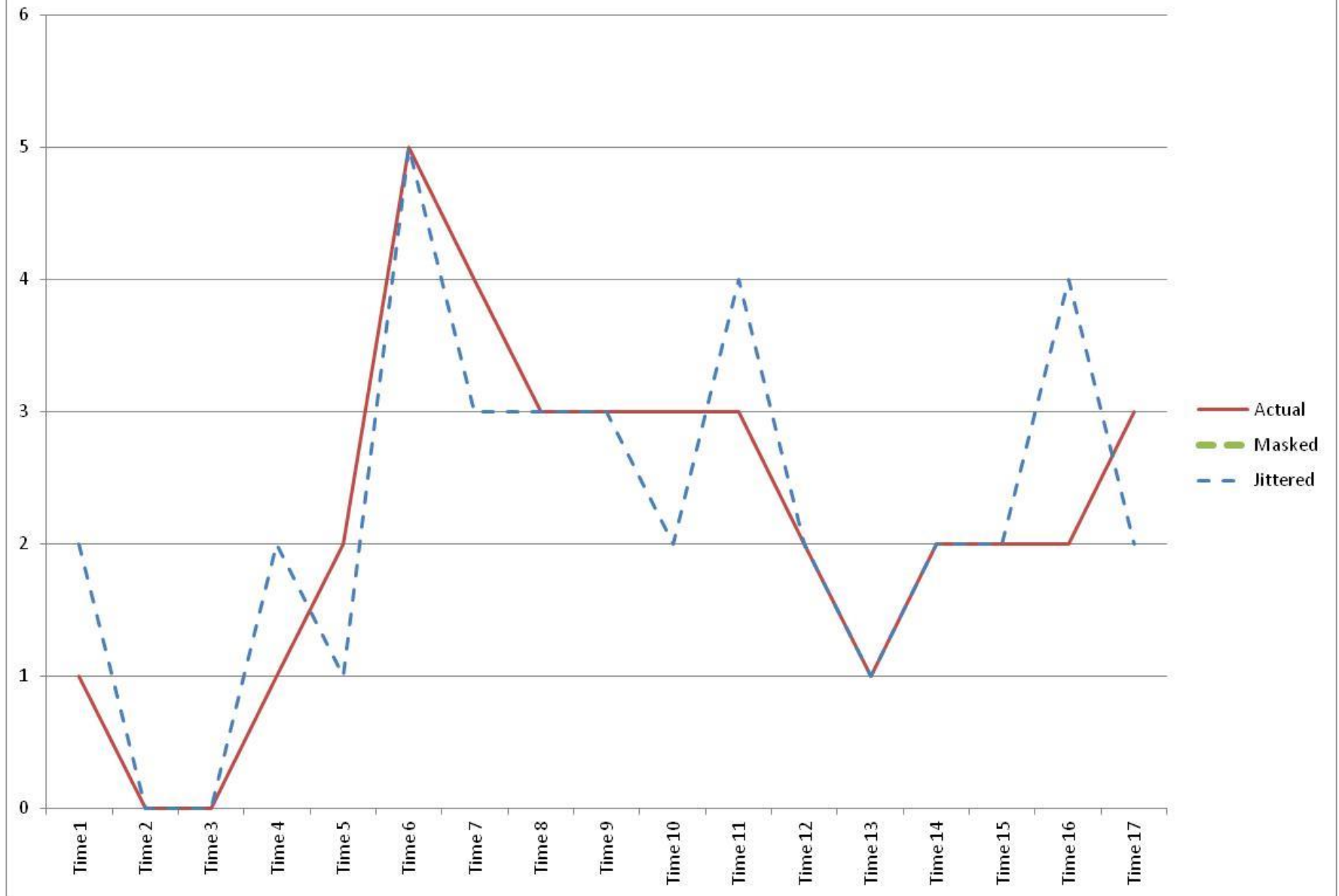
# Sample County 1--Number Authorized for Psychotropic Medications



## Sample County 1--Number Authorized for Psychotropic Medications



# Sample County 1--Number Authorized for Psychotropic Medications

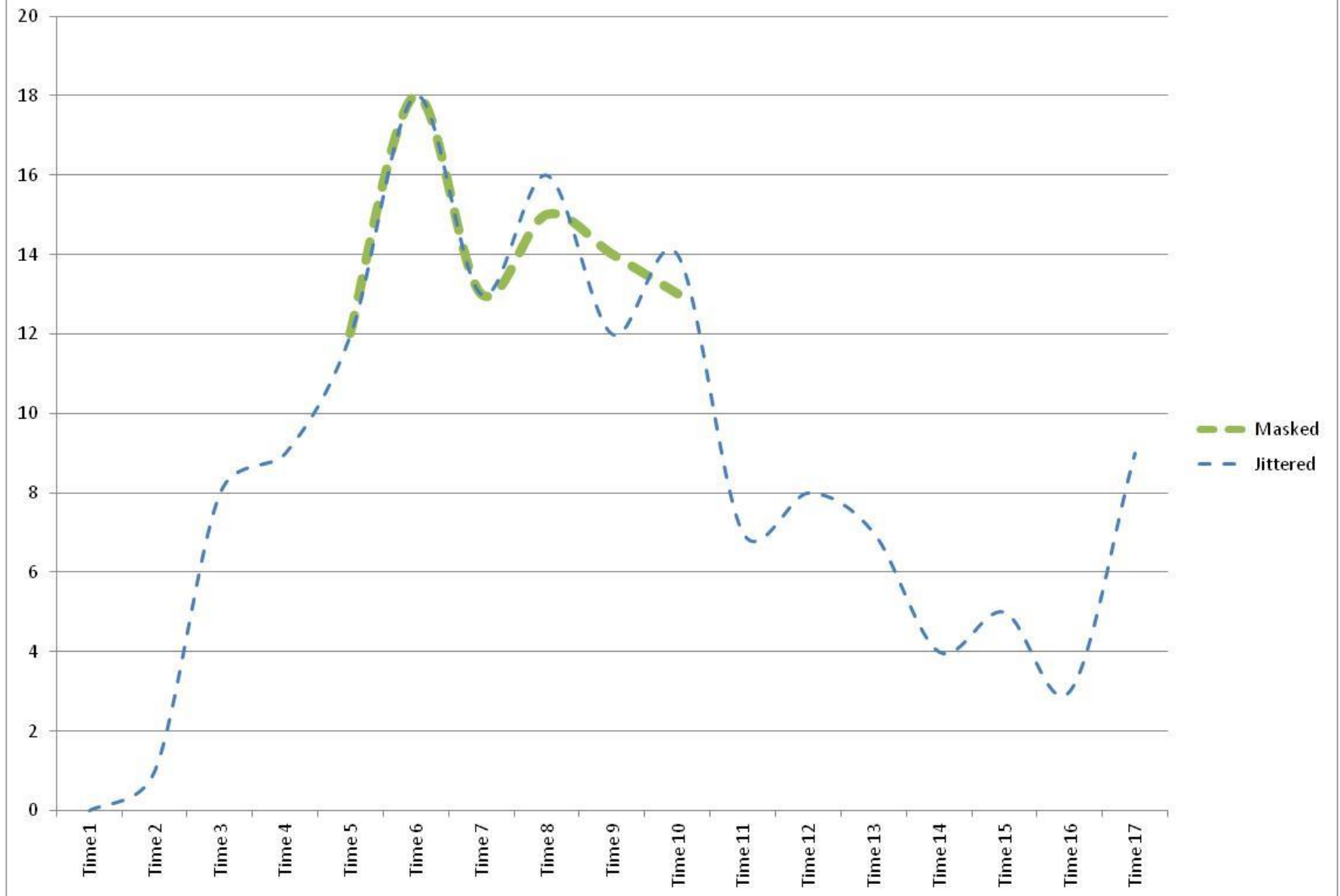


## Sample County 2--Number Authorized for Psychotropic Medications

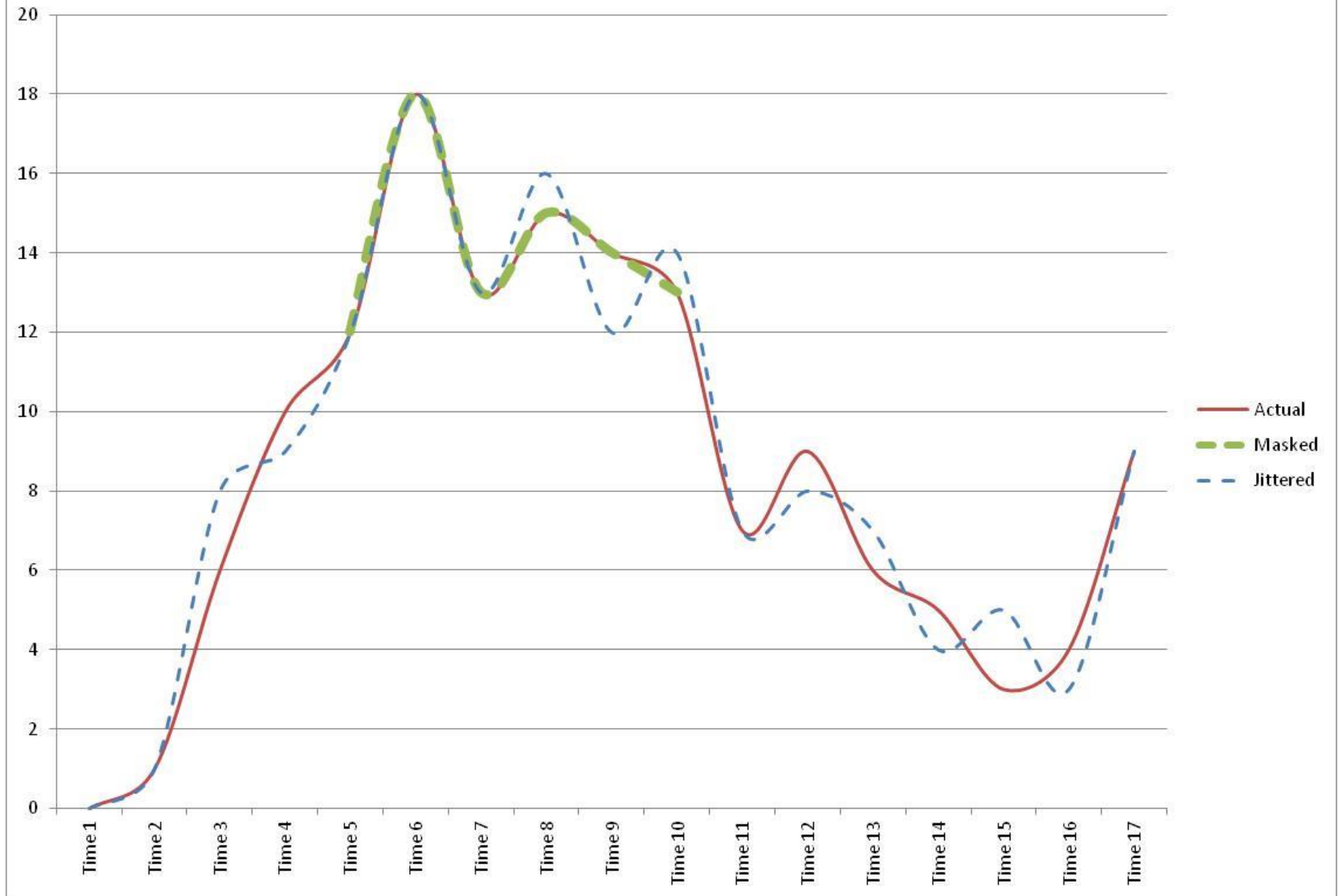




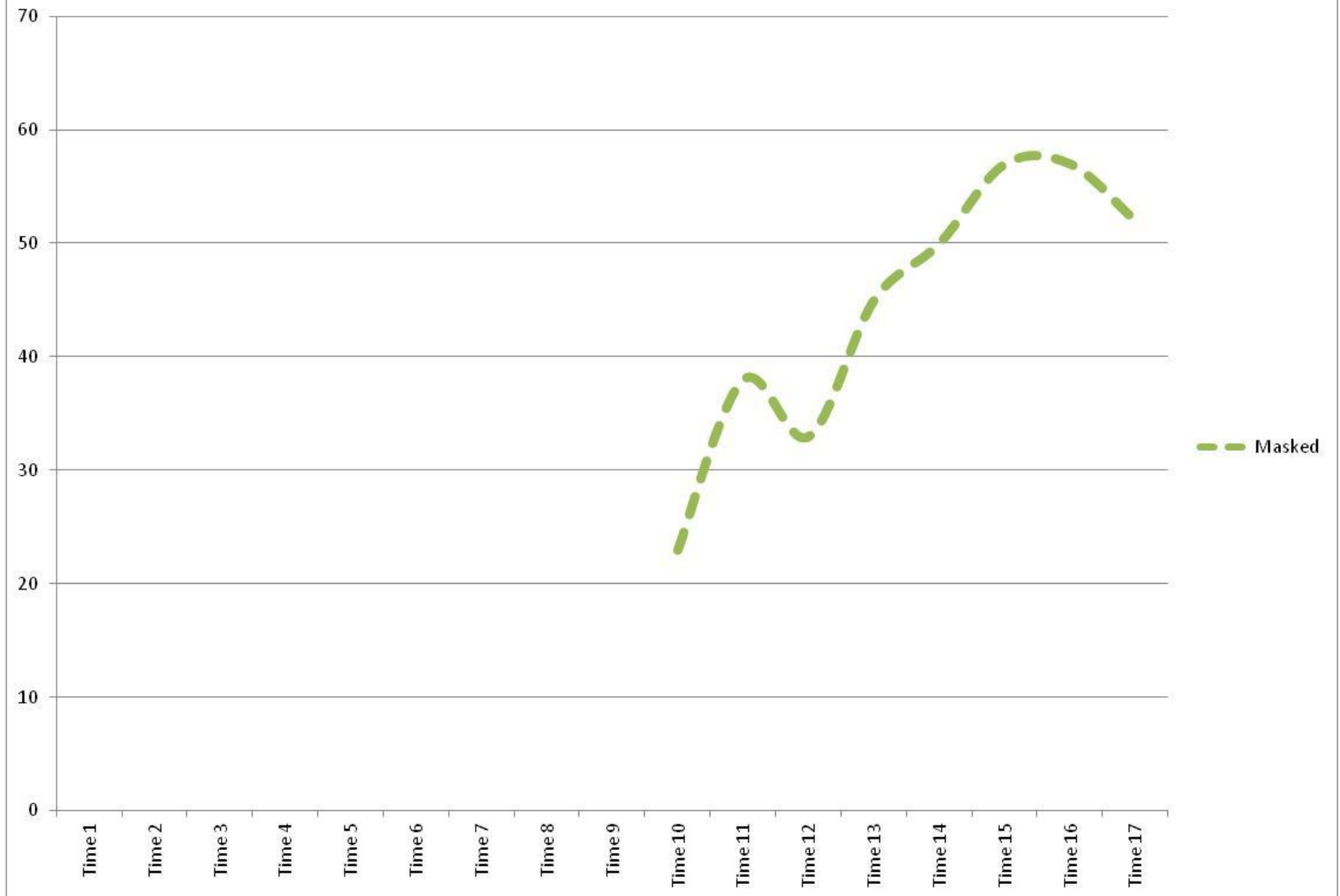
## Sample County 2--Number Authorized for Psychotropic Medications



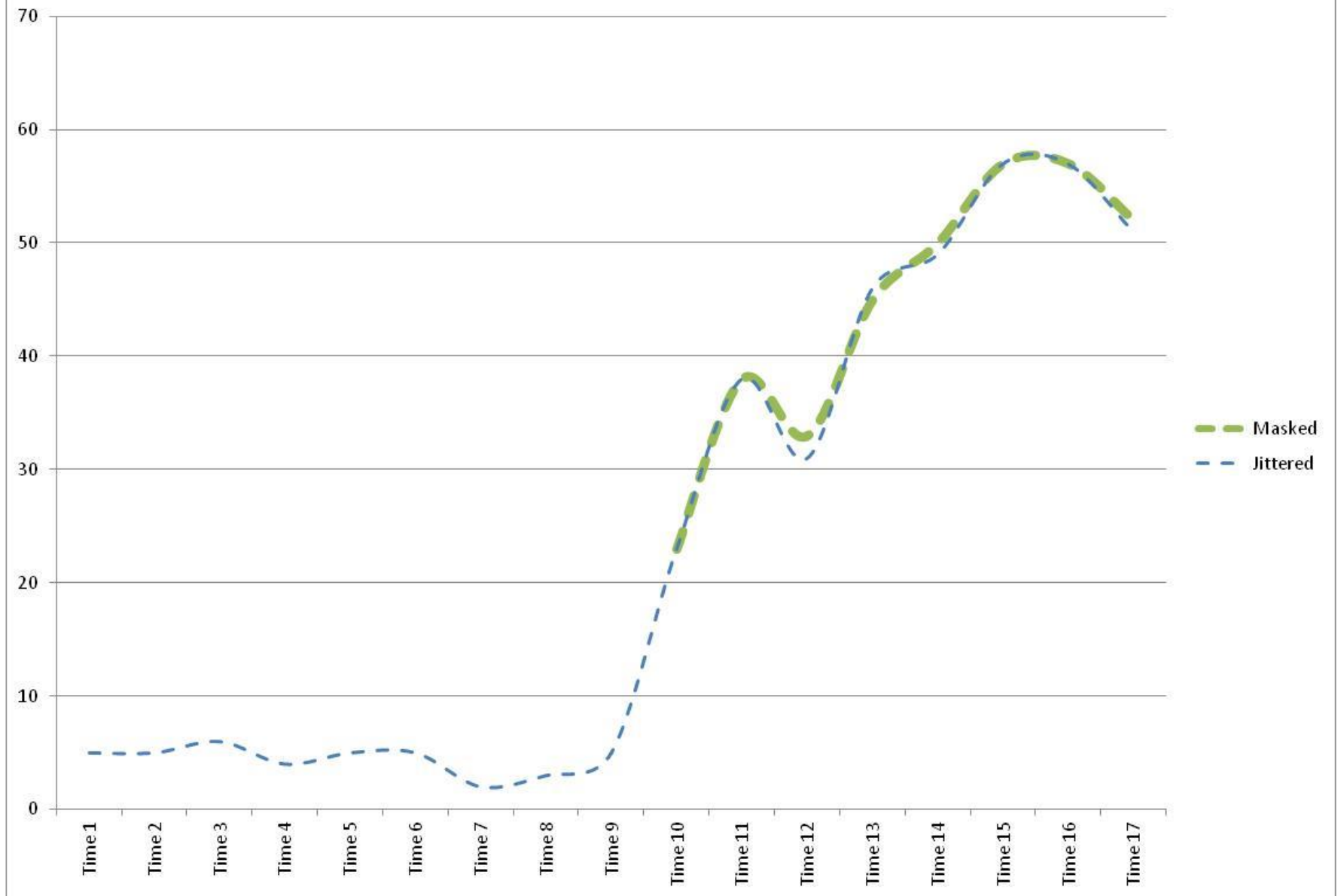
## Sample County 2--Number Authorized for Psychotropic Medications



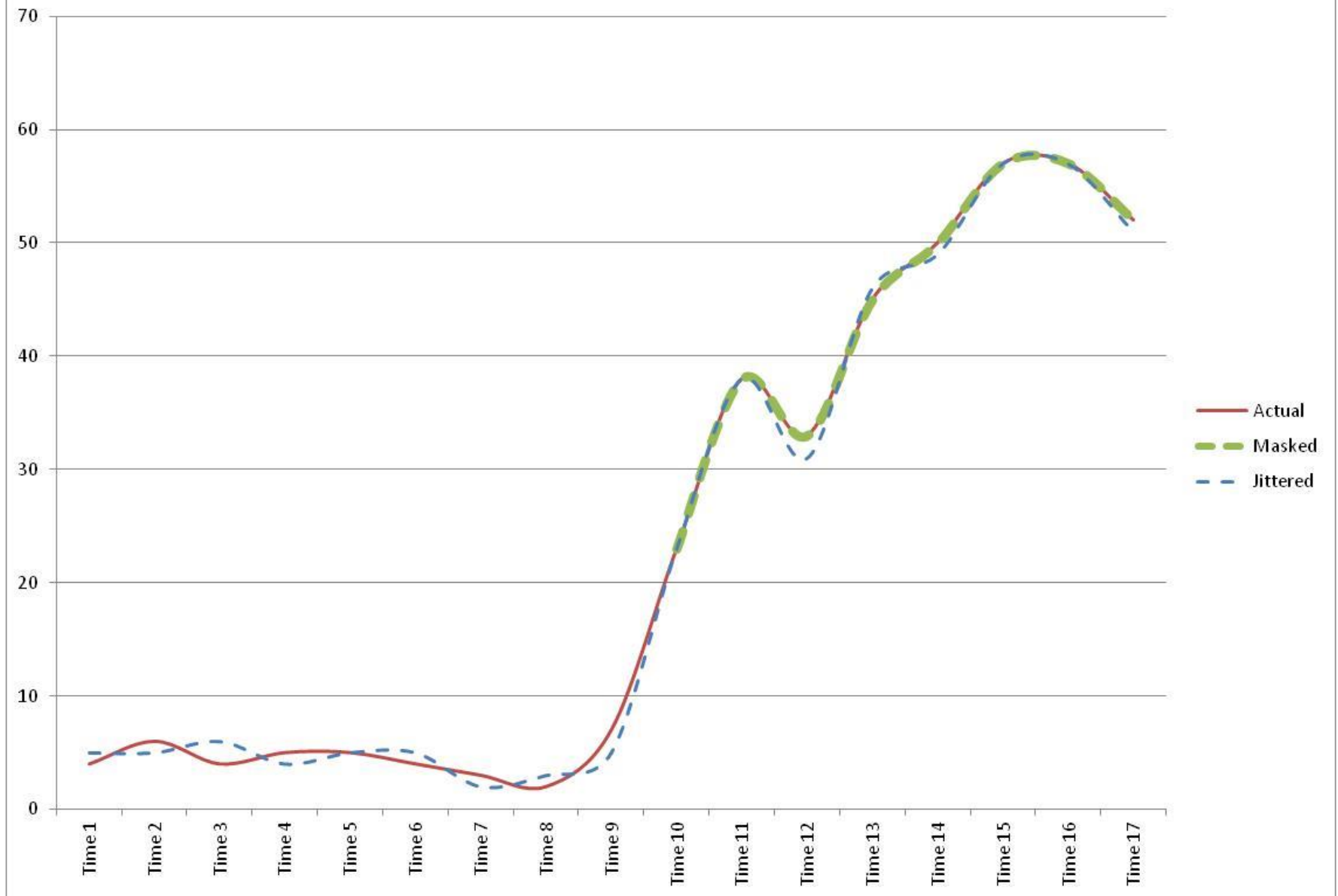
# Sample County 3--Number Authorized for Psychotropic Medications



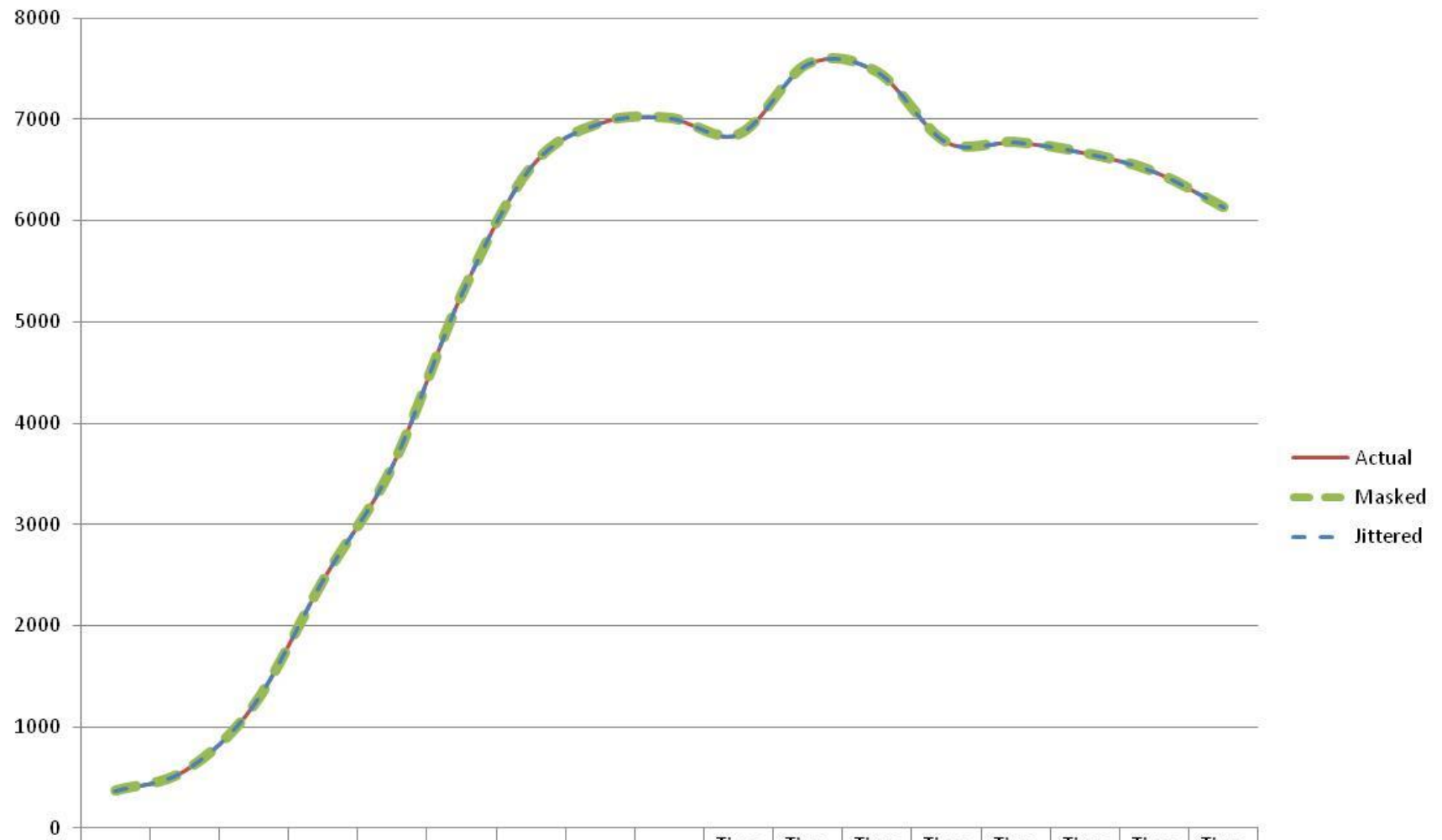
## Sample County 3--Number Authorized for Psychotropic Medications



# Sample County 3--Number Authorized for Psychotropic Medications



# California-Number Authorized for Psychotropic Medications



	Time 1	Time 2	Time 3	Time 4	Time 5	Time 6	Time 7	Time 8	Time 9	Time 10	Time 11	Time 12	Time 13	Time 14	Time 15	Time 16	Time 17
Actual	371	568	1,218	2,451	3,578	5,273	6,529	6,965	7,015	6,853	7,550	7,469	6,776	6,776	6,669	6,485	6,134
Masked	371	568	1,218	2,451	3,578	5,273	6,529	6,965	7,015	6,853	7,550	7,469	6,776	6,776	6,669	6,485	6,134
Jittered	370	568	1,217	2,451	3,578	5,271	6,530	6,965	7,015	6,854	7,550	7,471	6,775	6,776	6,669	6,484	6,135

# ***Questions?***

***510.290.6779***

[dwebster@berkeley.edu](mailto:dwebster@berkeley.edu)