Appendix 1

Materials and methods 1: Willingness-to-travel model

NHTS data is classified by modes of transportation (e.g., walk, air, bike, and privately operated vehicle) and purpose (e.g., work, school, family, and social). We analyzed national-level person trips by privately operated vehicle (POV) across all purposes. This included approximately 330,000 person trips (83% of all person trips in the database), totaling 3.3 million miles, including approximately 30,000 person trips originating in Texas.

We use the following notation:

Index set

$j \in J$ index set for distance bins from NHTS database

Data

d_j distance in miles of bin j

PT_j NHTS person trips from *j*-th distance bin

Model

$\hat{P}(d)$ fraction of target population willing to travel at least *d* miles:

$$\hat{P}(d) = \begin{cases} K_1 \exp(-\alpha_1 d^{\beta_1}) & \text{if } d < t \\ K_2 \exp(-\alpha_2 d^{\beta_2}) & \text{if } d \geq t, \end{cases} \tag{1}$$

where all model parameters (*K1*, *K2*, *a1*, *a2*, *b1*, *b2*, and *t*) are positive. Model (1) indicates that as the required travel distance grows, the fraction of the population willing to travel that distance drops. We use a piecewise model that allows for different coefficients below and above a distance threshold *t*, to capture the idea that rural populations have different behavior, with a greater willingness to travel. We fix *t* = 5 miles based on a preliminary analysis of the data we describe below.

Each record in the so-called "Travel Day" file in the NHTS database (*1*) represents one person trip, where a person trip is the most basic measure of personal travel. As an example, two household members traveling together in one POV counts as two person trips. The NHTS uses the following distance bins when recording person trips in miles: less than 1, 1-2, 2-3, 3-4, 4-5, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-74, 75-99, and 100 miles or greater. We use data for the first 16 of these 17 bins; i.e., we drop trips in excess of 100 miles and we use index set $J = \{1, 2, ...16\}$.

We denote the total number of person trips that fall in bin j by PT_j. In this way, we have empirical values for the fraction of the population willing to travel a distance of at least d_j miles for $j \in J$ by:

$$P(d_j) = \frac{\sum_{i \ge j} PT_i}{\sum_{i \ge 1} PT_i} \tag{2}$$

The denominator on the right-hand side of equation (2) is the sum of all person trips by POV for all bins included in set J. The numerator shrinks as the distance indexed by bin j grows.

Using a least-squares fit, we obtain the following model:

$$\hat{P}(d) = \begin{cases} \exp(-0.109 d^{1.184}) & \text{if } d < 5 \text{ (miles)} \\ 1.479 \exp(-0.4255 d^{0.6025}) & \text{if } d \ge 5. \end{cases} \tag{3}$$

A plot of the model in equation (3), along with NHTS data points, is given in Figure S1. The adjusted R^2 value for the fit exceeds 0.99 on both portions of the fit (i.e., both below and above the threshold distance of five miles) suggesting a good fit.

Two options for the target population include the entire population in Texas or the underinsured population. Model (3) provides input for our optimization model when our target population is the entire population of Texas. We perform a second fit for the underinsured population using NHTS data on household income associated with each person trip. We label a person trip as corresponding to an underinsured household if the income is $20,000 or less. The rationale for this follows. The average U.S. household size in 2009 (the year of the NHTS survey) was 2.63 (*2*). The 2009 Health and Human Services poverty guidelines indicate the corresponding annual income lies between $14,570 and $18,310 (*3*). NHTS income information is available in bins of $10,000, and. The $20,000 household income roughly corresponds to 138% of the Federal Poverty Level (FPL) for Texas. According to the Henry J. Kaiser Family Foundation, which provides state-level health insurance statistics for populations under age 65 (*4*), 51% of the non-elderly uninsured population in Texas falls below the 138% FPL threshold. Overall, 27% of the non-elderly population of Texas is uninsured, with 43% and 19% uninsured below and above the 138% FPL threshold, respectively. We believe that the <$20,000 income population provides a reasonable approximation for the geographic distribution of the underinsured population, and the decision to use this proxy was made in consultation with Texas DSHS officials involved in the 2009 H1N1 antiviral deployment and planning for future pandemics.

We proceed as before in carrying out a least-squares fit, but restrict attention to person trips reported from a household income of less than $20,000. Performing the least-squares fit we obtain:

$$\hat{P}_{at\_risk}(d) = \begin{cases} \exp(-0.1022 d^{1.21}) & \text{if } d < 5 \text{ (miles)} \\ 1.529 \exp(-0.433 d^{0.60}) & \text{if } d \ge 5. \end{cases} \tag{4}$$

As we mention in the main text, the results suggest that, relative to the entire population, the underinsured population has a very slightly greater (less than 1%) willingness to travel. The adjusted R^2 values for the fit again exceed 0.99, suggesting a good fit to the data. Next, we use these willingness-to-travel models in our optimization model.

Materials and methods 2: Optimization model

We formulate an optimization model that selects ZIP codes in which pharmacies should receive SNS and state cache antivirals to maximize access to antivirals for the user-selected target population (underinsured or entire population). Our model takes the form of a so-called facility location model from the operations research literature; see, e.g., Daskin (*5*). The model indexes ZIP codes by $i \in I$ and takes as input the target population in each ZIP code, denoted pop_i. We index a subset of ZIP codes that contain relevant pharmacies (as constrained by the user's choice of pharmacy chains and independent pharmacies) by $j \in J$. Using the centroid of each ZIP code we approximate travel distances between ZIP codes using great-circle distances. For travel within a ZIP code, we approximate distances assuming ZIP codes are squares giving $d_{ii} = \frac{\sqrt{(area_i)}}{2}$. To prevent any anomalies, we further take $d_{ij} = \max(d_{ii}, d_{ij}) \forall i \neq j$. Using these distances in the willingness-to-travel model of equation (3) or (4) we obtain the fraction of the target population of ZIP code *i* willing to travel to ZIP code *j* to obtain an antiviral dose, denoted p_{ij}. We summarize the model's notation and formulation as follows:

Indices and sets

$i \in I$ locations of customers [ZIP codes]

$j \in J$ ZIP codes containing candidate pharmacies that can be selected [subset of all ZIP codes]

Data

pop_i target population of customers in ZIP code i

$p_{i,j}$ fraction of target population in ZIP code i willing to go to ZIP code j to obtain an antiviral

b bound on number of ZIP codes selected

Decision variables

x_j if we select a pharmacy in ZIP code j and 0 otherwise

$y_{i,j}$ if ZIP code *j* contains ZIP code *i*'s most accessible pharmacy and 0 otherwise

Formulation:

$$\max_{x,y} \sum_{i \in I} \sum_{j \in J} pop_i \, p_{i,j} \, y_{i,j} \tag{5a}$$

$$\sum_{j \in J} x_j = b \tag{5b}$$

$$\sum_{j \in J} y_{i,j} = 1, i \in I \tag{5c}$$

$$y_{i,j} \le x_j \;\; i \in I, j \in J \tag{5d}$$

$$x_j \in \{0,1\}, \; j \in J \tag{5e}$$

$$y_{i,j} \in \{0,1\}, i \in I, j \in J. \tag{5f}$$

The objective function in (5a) specifies the expected value of the target population that has access to pharmacies with antivirals, and we seek to maximize that value. Constraint (5b) allows us to select pharmacies in at most *b* ZIP codes. The model assumes that the population at ZIP code *i* will choose to travel to the closest ZIP code that has a pharmacy with antivirals, as specified by constraint (5c), the objective function in (5a), and the fact that the willingness-to-travel function shrinks as distance grows. Constraint (5d) allows customers from ZIP code *i* to select a pharmacy in ZIP code *j* (i.e., $y_{i,j}=1$) only if ZIP code *j* has a pharmacy with antivirals ($x_j=1$). Constraints (5e) and (5f) require the pharmacy-selection decisions (x_j) and the access decisions ($y_{i,j}$) to be binary.

The meaning of the $y_{i,j}$ $y_{i,j}$ variables is subtle. For an individual located in ZIP code *i*, the $y_{i,j}$ variables answer the question: If you were to go to a pharmacy, which pharmacy would it be? In other words, $y_{i,j}$ determines, among all the selected ZIP codes, the one with the largest value of pi; j. According to the model, all of the target population in ZIP code *i* that chooses to go to a pharmacy will then go to that closest (i.e., most "accessible") pharmacy. Examining the objective function, and knowing from constraint (5c) that $\sum_{j \in J} y_{i,j} = 1$, we see that the variables have the effect of selecting the correct $p_{i,j}$ value for each ZIP code *i*.

Strictly speaking, model (5) takes as input the number of ZIP codes in which pharmacies can be selected, *b*. The model does not account for the number of antiviral doses DSHS has available, nor does it account for the number shipped to each pharmacy. Rather, we use the value of *b* as a surrogate for these quantities. The web-based tool solves model (5) for a family of values of *b* and displays an efficient frontier that depicts the trade-off between the expected target population with access to antivirals and the number of selected ZIP codes. And, the tool allows the user to select the solution for a particular value of *b* for further analysis, which we illustrate below.

To solve the mixed-integer programming model (5) we use CPLEX (*6*) called from GAMS (*7*). We maintain a tractable optimization model by eliminating variables $y_{i,j}$ for (*i, j*) pairs that have little effect. If travel from ZIP code *i* to *j* as predicted by the willingness-to-travel model is less than 5% we fix $y_{i,j} = 0$ and eliminate the corresponding variable from model (5). As a result of this, the population of some ZIP codes has no access to antivirals. However, we run a post-processing step and compute, the most accessible ZIP code (as predicted by the willingness-to-travel model) among those selected by decision *x*, and this is the accessibility.

For the hybrid optimization procedure, we run the optimization model sequentially. First we maximize access to the small ZIP codes alone with a fraction (75%) of *b*. Then we solve a second optimization which forces access in small codes to be at least 95% of that found in the first stage, with b ZIP codes, and we simultaneously maximize access using all ZIP codes.

Appendix 2

Here we present results and figures, previously presented only for the underinsured population, for the entire population of Texas. Again, small ZIP codes are those containing fewer than 1,000 underinsured people. These ZIP codes contain about 12% of the entire population of Texas (approximately 3 million people).

With 723 ZIP codes chosen by DSHS to provide access in 2009 (for details see main text) we estimate this network provides antiviral access for 87.9% of the statewide total population (Figure S3). By comparison, optimization over all possible pharmacy chains produces networks expected to achieve comparable access using only 606 ZIP codes, increase access to 90.9% with 723 ZIP codes, and reach a maximum access of 94.8% with all 1,023 ZIP codes. However, again the actual Texas 2009 distribution network and the corresponding optimized network (with 723 ZIP codes), are estimated to achieve only 53.9% and 55.2% access in small ZIP codes, respectively.

With 723 dispensing points, the hybrid method with *P* = 75% of dispensing points allocated to small ZIP codes produces networks that are expected to achieve 69% access in the small ZIP codes and 89.6% overall. For comparison, the highest possible access (when all pharmacies in the state dispense antivirals) is estimated to reach 72.6% and 94.8% in the two populations, respectively.

The major chains fail to reach the populations in the small ZIP codes, even with under the hybrid optimization that explicitly targets these hard-to-reach populations (Figure S3a). Broad statewide coverage is achieved by few major chains. For example, Walgreens alone is expected to achieve approximately 78% coverage if it dispenses in all of its 490 ZIP codes, with CVS and Walmart following closely behind (located in 422 and 372 ZIP codes, respectively). Independent pharmacies are still essential to bridging this gap in coverage. The maximum access achieved by a two chain combination in small ZIP codes is only 46.5%, with Brookshire and Walmart (Figure S3).

References

[1] U S Department of Transportation. Data Extraction Tool and Visualization Prototypes; 2009.

[Online; accessed August 30, 2013]. http://nhts.ornl.gov/det/.

[2] Nasser HE, Overberg P. After 50 years of decline, household size is growing; 2011.

[Online; accessed August 30, 2013]. http://usatoday30.usatoday.com/news/nation/census/2011-5-04-Census-Households-Demographics_n.htm.

[3] U S Department of Health and Human Services. The 2009 HHS Poverty Guidelines: One Version of the U.S. Federal Poverty Measure.

[Online; accessed August 30, 2013]. http://aspe.hhs.gov/poverty/09poverty.shtml.

[4] The Henry J. Kaiser Family Foundation.

[Online; accessed October 15, 2014]. http://kff.org/state-category/health-coverage-uninsured/?state=TX.

[5] Daskin MS. Network and Discrete Location: Models, Algorithms, and Applications. New York: John Wiley and Sons, Inc.; 2013.

[6] IBM. IBM ILOG CPLEX Optimization Studio CPLEX User's Manual,Version12, Release 4.

[Online; accessed August 15, 2013].
http://pic.dhe.ibm.com/infocenter/cosinfoc/v12r4/topic/ilog.odms.studio.help/pdf/usrcplex.pdf.

[7] Rosenthal RE. GAMS- A User's Guide ; 2013. [Online; accessed December 25, 2013]. http://www.gams.com/dd/docs/bigdocs/GAMSUsersGuide.pdf.

Table S1. Expected access (to the total population) provided by the Texas' 2009 H1N1 antiviral distribution network, a comparable size optimized network, and a full network containing all Texas pharmacies.

	Texas 2009 network	Optimized network[a]	All pharmacies
Small ZIP codes	53.9%	69%	72.6%
Statewide	87.9%	89.6%	94.8%
No. ZIP codes used	723	723	1,023

[a] Initially, we optimized 75% of dispensing points (542) to maximize access solely in small ZIP codes, and recorded the access achieved. Then, we optimized all 723 dispensing points to

maximize statewide access, while constraining the solution to achieve at least 95% of the small ZIP code access achieved in the initial optimization.

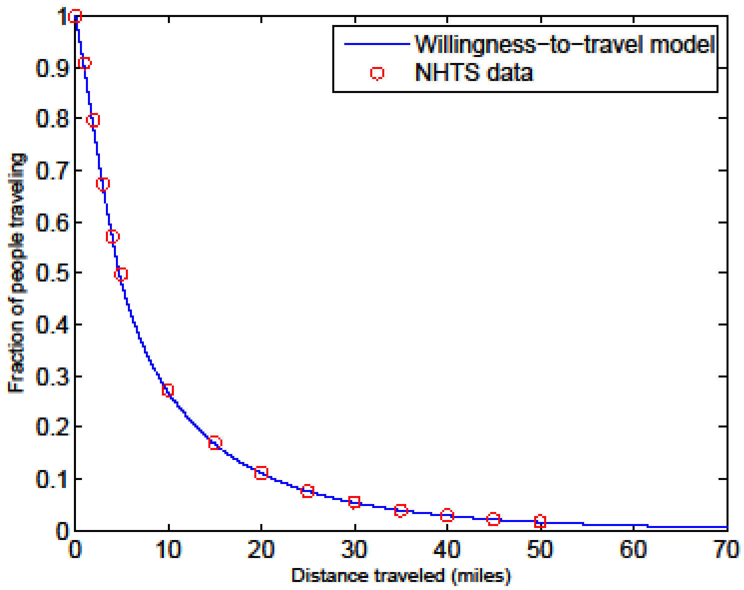


Figure S1. Willingness-to-travel curve given by equation (3) fit to NHTS data on POV travel for the entire U.S. population.

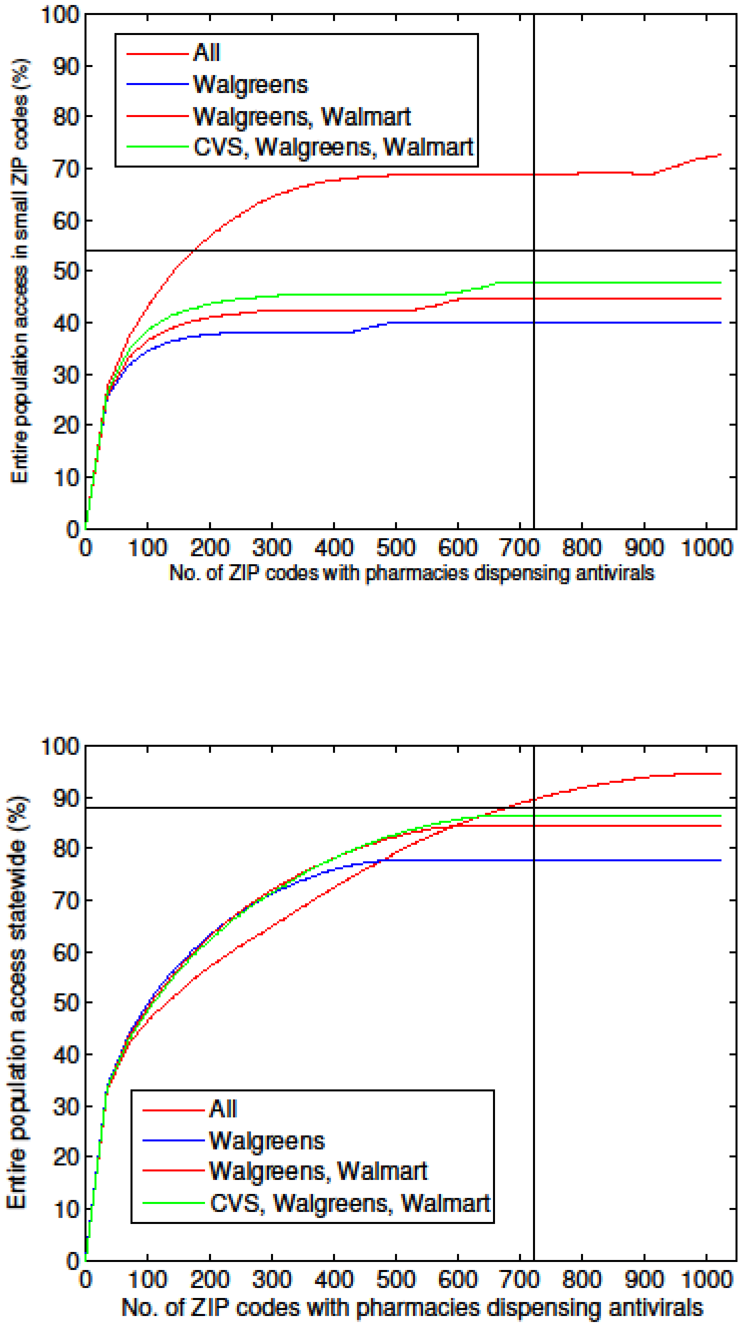
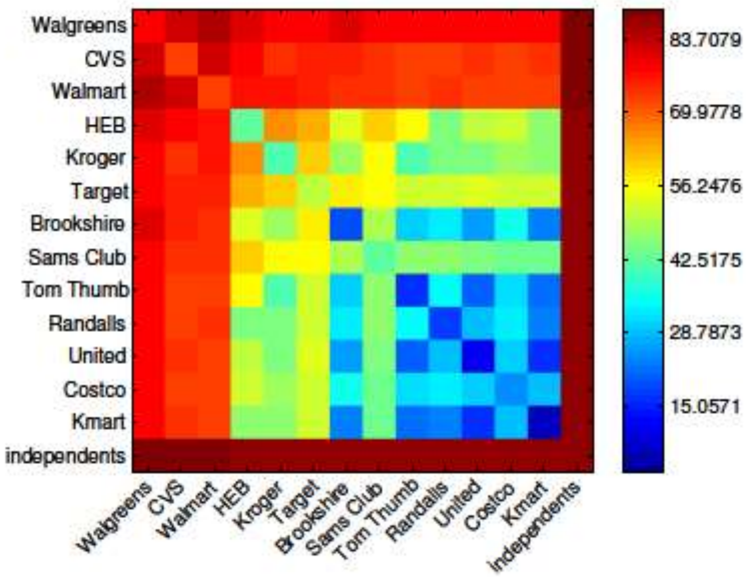


Figure S2. Antiviral access in total population achieved by the Texas antiviral distribution network during the 2009 H1N1 pandemic and by optimized antiviral distribution networks, for both (a) small ZIP codes and (b) statewide. Access is the expected fraction of the total population willing to travel to the nearest dispensing pharmacy to obtain antivirals. The black vertical and horizontal lines indicate the number of ZIP codes that participated in Texas' 2009 distribution network, and the estimated access achieved. For each network size (number of dispensing ZIP codes), a hybrid optimization was performed to maximize coverage in both small ZIP codes and overall (see text for details). Color indicates which combination of 13 major pharmacy chains plus independents were considered in the optimization. For a distribution network of size 723 (comparable to the Texas 2009 H1N1 antiviral distribution), the best performing single chain (Walgreens), two chain combination (Walgreens and Walmart), and three chain combination (Walgreens, Walmart, and CVS) provide near optimal coverage statewide, but critically underserve the smallest ZIP codes.



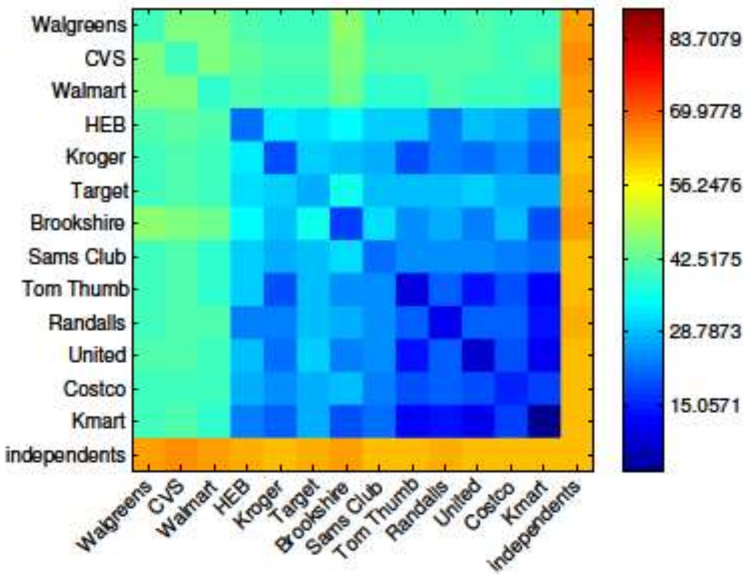


Figure S3. Antiviral access in total populations for single chain and two chain pharmacy distribution networks. Each network contains a maximum of 723 dispensing points, and was designed using a hybrid optimization that maximizes coverage in both small ZIP codes and overall (see text for details). Color indicates the expected percentage of the total population willing to travel to dispensing pharmacies to obtain antivirals (a) statewide and in (b) small ZIP codes. Diagonal entries correspond to single chain networks.

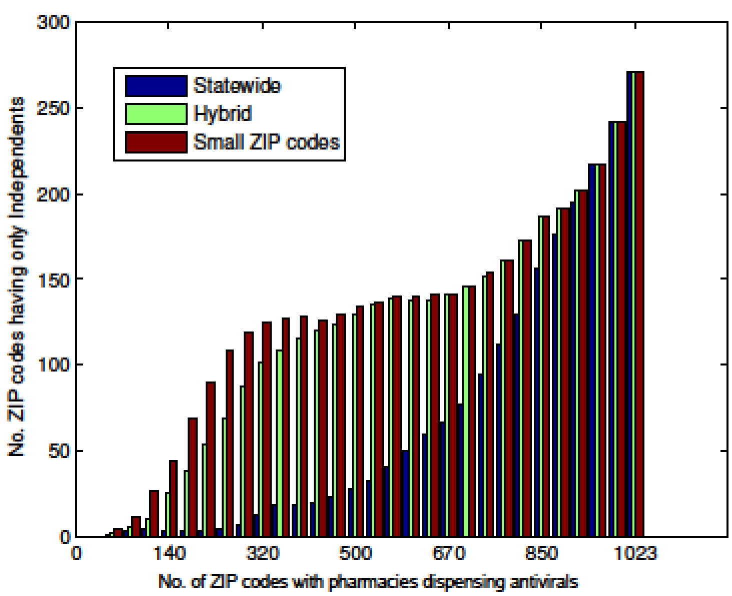


Figure S4. The number of sites in the antiviral distribution network containing only independent pharmacies (no major chains) when optimizing for the total population in small ZIP codes, statewide, or both (hybrid).

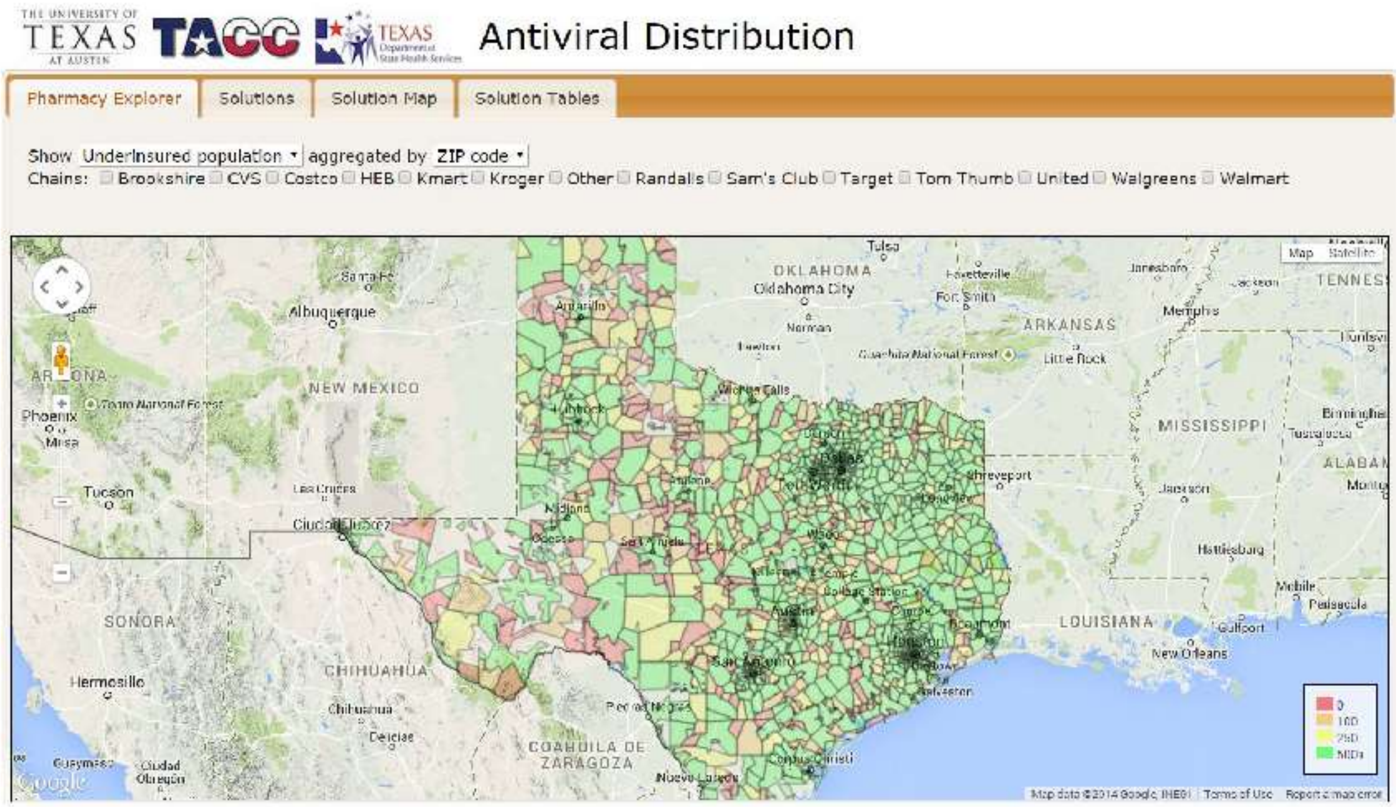


Figure S5. Screenshot of Texas Antiviral Distribution tool, showing underinsured population of each ZIP code.