Institute for International Economic Policy Working Paper Series
Elliott School of International Affairs
The George Washington University

Can Differences Deceive? The Case of "Foreclosure Externalities"

IIEP-WP-2017-29

Anthony Yezer
George Washington University

Yishen Liu
George Washington University

November 2017

Institute for International Economic Policy
1957 E St. NW, Suite 502
Voice: (202) 994-5320
Fax: (202) 994-5477
Email: **iiep@gwu.edu**
Web: **www.gwu.edu/~iiep**

Can Differences Deceive? The Case of "Foreclosure Externalities"

Yishen Liu* Anthony Yezer†

July 19, 2017

Abstract

Foreclosure externalities, in which recent foreclosures proximate to a housing unit depress its sales price, are well accepted in the literature. These papers use a geographic differencing strategy to eliminate the problem of selection into treatment. They also assume that the partial and total derivatives of the outcome (house value) with respect to the treatment (foreclosure) are constant and equal. This paper relaxes these assumptions producing very different results. These findings likely generalize to a larger body of research where differencing often in the form of regression discontinuity, propensity score matching, or synthetic controls is used to achieve identification while assuming total and partial derivatives of the outcome with respect to the treatment are constant and equal.

Keywords: Foreclosure; Specification error; Loan-to-value ratio; Externalities.

JEL Classification Numbers: R23, R30, R31.

*Economist, Fannie Mae.

†Department of Economics, George Washington University, address correspondence to yezer@gwu.edu.

1 Introduction

The high rate of foreclosure on single family mortgages in the period following 2006 provided a promising opportunity to investigate the possibility that the foreclosure process generates local externalities. Foreclosure rates varied substantially both over time and across areas within a given housing market. Many papers found a negative relation between recent foreclosures and the current sales price of housing proximate to those foreclosures. These papers recognize that local unobserved heterogeneity in housing markets may generate the "treatment" of foreclosure and they adopt clever differencing techniques to eliminate the effects of this correlation on the relation between foreclosure and house price. The issues raised in this paper have nothing to do with identification problems of selection into treatment that motivate or mandate differencing to identify causal effects.[1]

Several theoretical mechanisms have been proposed that could connect foreclosed properties to local house prices. Some of these are real effects. The most direct relation is a real physical externality due to decreased maintenance of the foreclosed property that makes it less attractive to neighbors. Gerardi et al. [16] have called this the "investment effect". To the extent that foreclosure results in eviction or at least forced migration, there might be social effects on a neighborhood. There is evidence that foreclosed properties are more likely to be purchased by investors changing the tenure mix of a neighborhood. If sales prices rise with owner occupancy rates, this "tenure mix effect" could lower prices.[2] To the extent that foreclosure sales are involuntary, they reflect an exogenous increase in local supply which might temporarily lower prices. Annenberg and Kung [1] have argued for this "supply effect" mechanism. Substantial evidence indicates that sale of REO by lenders occurs at depressed prices, likely reflecting the failure to market the property in an optimal fashion

[1] In the case of the research results presented in this paper, bias due to selection into treatment is not a concern for two reasons. First, the identification strategy is based on the approach used in previous papers. Second, and most important, the direction of any simultaneous equation bias is negative, i.e. in favor of the alternative hypothesis. Given that the final results in this paper reject that alternative, there is no danger that this rejection is due to type II error based on selection into treatment.

[2] See Coulson and Li [10] for an attempt to deal with the difficult problems of making this inference about owner occupancy rates and house prices.

and/or damage done to the unit in the foreclosure process. To result in an externality, there must be a mechanism to translate the lower prices of foreclosure sales to the sales of non-distressed properties. This could occur if there is an "appraisal effect" in which appraisers and buyers fail to distinguish effects of property condition on foreclosure sales prices and lower appraisals of non-distressed properties that are well maintained. In sum, previous papers have identified several theoretical mechanisms that could produce empirical results indicating a negative externality associated with foreclosure.

Based on both a review of the literature and new results reported here, it appears that the investment effect is very important. Investment effects could also be produced because owners are delinquent or lack equity in the property. Of course both these conditions are prerequisites for foreclosure. However, transfer of title in foreclosure ends the conditions of delinquency and negative equity putting the housing unit on the path to sale, recapitalization, and investment. This raises the possibility that foreclosure is actually the cure for what should be called delinquency and negative equity externalities. A properly specified model of the relation between sale prices of housing and proximity to properties in all stages of financial distress is needed to test for the nature of these externalities.

Past tests of the association between foreclosure and house values have acknowledged difficulties created by selection into treatment. Falling house prices can produce negative equity and induce households to default, which ultimately results in foreclosures. This possibility has been widely recognized. The standard response in the literature has been to use spatial and temporal disaggregation to eliminate the possibility that localized house price effects could have caused nearby foreclosures. By saturating the estimating equation with time and locational dummies at the sub-neighborhood (census tract or smaller) level and lagging foreclosures, researchers argue that they have removed any correlation between foreclosure and differences in price trends. The techniques effectively pair units co-located in the same sub-neighborhood which vary in their proximity to previously foreclosed units in that same sub-neighborhood during a particular time interval.

This paper does not dispute the assumption that such differencing and lagging removes the effects of selection into treatment and identifies a causal relation. However the estimation technique still involves strong assumptions and it obscures the role of alternative variables in the process being studied. First, stochastic specifications have the implicit assumption that the relation between treatment (foreclosure) and outcome (housing price) is additive independent. In complex systems, the assumption that the derivative of the outcome with respect to a treatment is constant may be difficult to justify. The standard response to those who question the assumption of additive independence is to say that, if there is heterogeneity in treatment effects, the estimates reflect an average treatment effect. This view suggests that a fully specified model capable of revealing details of the underlying mechanism responsible for the statistical association is not essential. The results presented here demonstrate that this is certainly not the case for foreclosure externalities where the relation between foreclosure counts and property values is not additive independent and this insight significantly changes the interpretation of empirical results that have appeared in the literature.

Second, even if the relation between house value or its logarithm and foreclosure is additive independent, proper specification requires that other variables that vary spatially at the sub-neighborhood level be considered, particularly if this variation could be correlated with the treatment. But, even if the omitted variables are relatively uncorrelated with the treatment and/or instruments uncorrelated with the omitted variables are found, omission of variables reduces understanding of the causes of variation in the outcome. This problem will be termed the "micro variation" omitted variable problem. Insofar as practicable given limits of data availability, variables that are important to micro variation of the outcome should be included in the model. In the case of foreclosure externalities, understanding the relation between foreclosure as a treatment and local house prices as an outcome is significantly modified as a result of extending the estimation results to include other variables whose micro variation is similar to foreclosure.

This paper identifies variables important to the relation between foreclosure and house

prices which illustrate the effects of, additive independence and micro variation of omitted variables. The econometric problems created when the underlying relation is not additive independent and/or important variables are omitted are well understood. The contribution here is to illustrate how radically the interpretation of empirical results can change as a result of recognizing and correcting for these problems in a well accepted differences in differences research design.

The variables in the foreclosure externality literature that create the additive independence problem are the number of housing units in the micro area surrounding each observed transaction and measures of the strength of creditors' remedies. Curiously these easily observed variables have been ignored or entered in an additive independent manner in all of the literature reviewed in the next section of this paper. It is certainly possible that a foreclosure is more consequential for valuation of a nearby property if it is one of 10 units proximate to the home being sold than if it is one of 100 or 1,000 units nearby. This hypothesis is easily tested empirically and it is validated in the empirical work later in this paper. Note that this is not a problem of micro variation although nearby housing unit counts may vary within small areas. The appropriate specification involves the ratio of foreclosed units to total units, a relation based on the quotient of the two variables rather than their sum.

Creditors' remedies determine the ease of foreclosure and the ability of lienholders to protect the condition of delinquent properties as they move through the foreclosure process. Cordell and Lambie-Hanson [8] document the substantial variation in foreclosure cost over time and across jurisdictions as the housing crisis deepened. There were two reasons for this. Statutory changes weakened remedies and the institutions of the market that administer foreclosure became congested by the volume of activity. The delay in foreclosure substantially lagged recovery in the housing market. This, of course, implies that the size of any foreclosure externality due to the investment effect should not be constant over space and time as has been implicitly assumed in additive independent estimates where specifications require that all foreclosures have an equivalent effect on house prices.

The variables exhibiting micro variation that are generally omitted in the literature are counts of both seriously delinquent mortgages and units with high current loan to value ratios (CLTV) proximate to the house being sold. Delinquent borrowers or those with high CLTV should rationally choose lower levels of maintenance that could create the same local price effects as the investment effect of foreclosure. Gerardi et al. [16] have added the spatial distribution of delinquencies to a foreclosure model and found that the effects of foreclosure appear to be reduced even if additive independence is assumed. More important, they find that serious delinquency depresses surrounding property values independent of foreclosure. Biswas and Davidoff [2] measure the spatial distribution of high CLTV properties around each property sale but use it as an instrument for foreclosure rather than as an independent determinant of property prices. These omitted variables are chosen strategically because both serious delinquency and high CLTV should be necessary conditions for foreclosure.[3]

This paper estimates a standard foreclosure externality model incorporating two variables entering multiplicatively, the inverse of the number of nearby housing units and foreclosure cost (as a measure of effective remedies), and two micro variation variables, proximate high CLTV and delinquent properties. The estimation results confirm that the proper specification of a foreclosure externality equation is the ratio of foreclosures to housing units, not the level of foreclosures. Perhaps most important, the absolute size of the relation between foreclosure and local house prices appears to vary directly with the difficulty of foreclosure. Foreclosure cost is the primary determinant of the effect of foreclosure on local house prices. Both of the two micro variation variables, the fraction of high CLTV units and seriously delinquent units have similar but smaller depressing effects on nearby housing prices given current policies toward creditors' remedies. Estimation results indicate that the relation between house prices and the ratio of foreclosures to housing stock is convex which is consistent

[3]Let V be property value, and postulate that $V = v(F, D, L)$ where F, D, and L are measures of nearby foreclosures, serious default, and high CLTV properties respectively. Even assuming additive independence of F, D, L, so that v_F is a constant, the total treatment effect of F is given by $\frac{dV}{dF} = v_F + v_D \frac{dD}{dF} + v_L \frac{dL}{dF}$. Given that foreclosure removes seriously defaulted and underwater owners, $\frac{dD}{dF}, \frac{dD}{dF} < 0$. Accordingly if, as seems plausible, v_D, $v_L < 0$, then $\frac{dV}{dF} > 0$ is possible even if $v_F < 0$.

with the investment effect and inconsistent with the sales effect.[4]

The view of foreclosure externalities implied by the approach and results in this paper is substantially different than the previous literature. Both seriously delinquent properties and high CLTV properties that are not in foreclosure produce negative external effects on local sales prices. The mechanism appears to be an investment effect. A lengthy foreclosure process increases these negative effects. But foreclosure also reduces the number of delinquent and high CLTV units. Ultimately foreclosure, even if delayed, cures the externality. Regulatory responses to the foreclosure externality literature have lengthened the foreclosure process and likely made the externality problem worse as foreclosure delays increase the number of delinquent and high CLTV units. Furthermore the specification and interpretation issues identified here may generalize to other empirical work on treatment effects where strong assumptions regarding additive independence and micro variation of omitted variables have not been justified. The arguments made here are similar to the recent critique of RCTs by Deaton and Cartwright [13].

2 Literature on Foreclosure and Nearby Property Value

Nature and diversity of the foreclosure process

The empirical literature uniformly reports the existence of foreclosure externalities even when there is substantial diversity in the nature of foreclosure. The foreclosure process in states that require judicial foreclosure must begin with court action in which the lienholder sues the borrower.[5] In states that allow statutory foreclosure the trustee acts. Notice must be filed as provided in the deed of trust but there is no need to sue the borrower. In either

[4]Some papers have attempted to test for non-linearity in the effects of number of foreclosures on sales price. The results have been rather mixed. Perkins et al. [23] report a threshold effect in which the first two foreclosures have no significant price effect so that the overall relation between foreclosures and sales price is concave. The shape of the relation between foreclosures and price has been identified as important for discriminating between the investment and supply effects.

[5]These actions are usually unopposed but they are court actions and result in court orders.

case, the borrower receives a "foreclosure complaint" documenting that there is a sufficient delinquency in scheduled payments to warrant foreclosure.[6] The borrower may respond to the complaint by curing the deficiency or mounting a defense which requires a trial. Even in states that require judicial foreclosure, most lienholder motions for summary judgement are granted. If the foreclosure process is not interrupted as discussed in some detail below, it results eventually in a "foreclosure judgement" followed by a "foreclosure auction" where title to the property is transferred. The empirical literature on foreclosure externalities uses either the initial foreclosure complaint or the transfer of title as an indicator of a property experiencing foreclosure. Clearly these are very different events. The former indicates serious delinquency and the latter failure to resolve the delinquency and transfer of title.

Borrowers often respond to the complaint by curing the delinquency. The first recourse is to the lienholder who may be willing to consider a loan modification, forbearance agreement, or payment plan. In some states the borrower may request a mortgage modification mediation hearing. The borrower may self-fund the cure, or refinance the mortgage. Any of these steps reestablishes the borrower's equitable right of redemption. Selling the property is another possibility. In cases of negative equity, refinancing is likely not possible and selling will not cure the entire debt to the lienholder.[7] However, a short sale can be arranged in which the borrower agrees to cooperate with efforts to sell the property and the lienholder agrees to accept the sale proceeds as full payment of principal, interest, and penalties outstanding. In the dataset used here, approximately half of the foreclosure notices do not result in sale of the property.

Other borrower responses are possible. A bankruptcy filing can substantially delay or

[6]In states with judicial foreclosure the lienholder must sue the borrower in court and obtain notice, or lis pendens, which is public information. Where statutory foreclosure is permitted, the trustee sends a notice of delinquency to the borrower who then has a period to cure the deficiency before the property goes to auction. Judicial actions are in the public record and statutory notice of intent to foreclosure must also be filed with the county recorder allowing local vendors to collect and publish this information which is available to investors, realtors, etc. The past standard of 90 day delinquency was extended to 180 days in 2014 by the Consumer Financial Protection Bureau.

[7]At this point negotiation between lienholder and borrower are likely heavily influenced by whether the state permits deficiency judgments against the borrower. Deficiency judgments are permitted in all three jurisdictions in the data set used here.

even reverse foreclosure. The borrower may also invoke the right of rescission under the Truth in Lending Act (TILA). Regardless of delinquency status, if the borrower informs the lender of the intent to rescind based on violation of TILA, this stops the foreclosure process and sets in motion a series of legal options that could permanently impair the lender's security interest in the collateral.[8]

If there is no interruption in the foreclosure process, i.e. the borrower does not act to reestablish the right of redemption or surrender the deed in lieu of foreclosure, an auction is scheduled. This is usually conducted by an officer of the court. The lender generally participates in this auction and, in the absence of other bidders, takes title to the property. Alternatively, title can be transferred to another individual or organization bidding higher than the lender. If the lender takes title, the property is then REO which is sold either immediately or after being held for some period of time.

In some states, borrowers retain a right of redemption even after disposition of the property at auction. This right of redemption may extend for as much as a year, clouds the title available at auction, and may delay final disposition of the property to a willing buyer. Because of the complications associated with this extended right of redemption, the areas selected for study here do not have this provision. In some states, eviction is an issue because the foreclosure judgment converts the borrower into a renter and eviction from rental units can be protected. In the end, eviction may be accelerated by payment of key money. These details are potentially important and have been embodied in the measure of foreclosure difficulty used here. There can be significant delays in the ultimate disposition of the property to a new owner with clear title and an incentive to maintain, and even upgrade the property. In sum, there are many reasons why the relation between foreclosure actions and condition of the real property collateral may vary across states due to substantial variation

[8]This three-year right to rescind under TILA does not extend to home purchase mortgages, i.e. primary purchase mortgages or piggy back seconds, or financing of property which is not a primary residence. Borrowers with a first and non-piggy back second from the same lender can exert an extra measure of influence with the lender if they rescind the second. Rescission brings the loan out of technical default until further legal action or negotiation determines the response to the claim of imperfect disclosure under TILA.

in creditors' rights and the performance of the court system.

As noted above, some empirical tests for foreclosure externalities count the foreclosure complaint or lis pendens in the case of judicial foreclosure while others define foreclosure as the transfer of title from the borrower sometime after the complaint, usually at the time of the foreclosure judgment. These two measures are correlated but certainly vary significantly across jurisdictions, as documented later in the data section. To the extent that the link between foreclosure and nearby property values is based on maintenance issues, transfer of title is a more consequential measure because, in almost all cases, it implies negative equity in the unit. There are two reasons for this. First, the borrower's failure to cure through refinancing or from internal funds implies high CLTV, likely exceeding unity. Second, borrowers not only lack an incentive to maintain units when title is being transferred completely and permanently, they may also damage the unit as they vacate. Borrowers who cure their delinquencies are indicating that they have equity which is worth preserving by curing, maintaining, and not damaging the property. Thus far, empirical studies where foreclosure is based on notice and those requiring evidence of property transfer have produced similar results, i.e. an increase in the number of foreclosures proximate to a unit tends to lower the sales price of the unit. Furthermore, studies using data from states requiring judicial foreclosure are similar in their finding of external effects on sales prices to those from states where statutory foreclosure is permitted. The uniformity of empirical measures of foreclosure externalities given the diversity in what is measured as a "foreclosure" suggests that the process generating lower prices in areas with more foreclosure proceedings needs more careful consideration.

Own-price effects of foreclosure sales

The foreclosure externality literature builds on research that has established stylized facts about the sales price of foreclosed properties. Foreclosed properties sell for less than similar surrounding units. Pennington-Cross [22] finds that prices of distressed properties appreciate

22% less than a price index of non-distressed sales. Clauretie and Daneshvary [7] report a smaller but still substantial foreclosure discount. Campbell et al. [4], not only find a 27% price discount for foreclosures, they argue that part of this is due to the forced nature of the sales process as lenders unload REO. They report price discounts for other situations, such as estate sales, which are grouped under the general category of forced sales. Sumell [26] finds even larger, perhaps 60%, price discounts in inner-city Cleveland. Recently Mian et al. [21] found that, during the 2007 to 2009 portion of the housing downturn, states permitting statutory foreclosure experienced far more foreclosure and much sharper price declines than nearby states that only allow judicial foreclosure. These effects abated by 2011 to 2013. Lambie-Hanson [20] reports that public complaints about properties undergoing foreclosure are 1.3 to 1.7 times as frequent as otherwise comparable properties. This is a small sample of a literature that documents the existence of a substantial foreclosure discount which arises from the deterioration in condition of housing units as they pass through the foreclosure process. Deterioration may increase with the volume of foreclosure activity and vary across housing markets. All this is consistent with the "investment effect" where lack of maintenance imposes a physical externality on nearby housing and the "appraisal effect" in which comparables include foreclosure sales at depressed prices.

Testing for the existence and size of foreclosure externalities

Testing for an empirical effect of nearby foreclosure on property values has produced a fairly uniform set of conclusions regardless of the definition of foreclosure, initial complaint versus judgment and property transfer. Furthermore, differences in the regulatory environment discussed above, including judicial versus statutory foreclosure, deficiency judgements, redemption, etc., do not alter findings. Even variation between measuring price effects using classic hedonic versus repeat sales price indexes still produces the same general negative relation between the number of recent foreclosures proximate to a housing sale and the price

at which the transaction occurs.[9] By 2010, the literature was large enough to support a review article, Frame [15].

In addition to the general measurement of foreclosure externalities, some studies have attempted to use special situations to test alternative hypotheses regarding the causes of foreclosure externalities. One result, established by Fisher et al. [14], is that foreclosure externalities are much less important in the multifamily housing stock. Specifically effects within a given structure are much larger than spillovers across structures. Whitaker and Fitzpatrick [28] find that vacant properties produce a negative external effect on surrounding property sales similar to that of foreclosed properties. Results in Cui and Walsh [11] suggest that vacancy and foreclosure are associated with higher crime rates.

Gerardi et al. [16] come closest to the tests for foreclosure externalities in this paper. In addition to foreclosure effects on local housing price, they add a measure of the number of delinquent properties and find that both foreclosure and serious delinquency, but not early delinquency, have negative effects on local sales price.[10] They also have some information on property condition for foreclosures and find that the negative foreclosure effect varies inversely with property conditions, suggesting an investment effect.

Groves and Rogers [17] provide strong evidence that foreclosure externalities are produced by the investment effect. In a natural experiment, they test for the size of foreclosure externalities in ordinary housing compared to housing located in residential communities where homeowners' associations have covenants in deeds that allow them to mandate exterior maintenance. They measure both the own-price and externality effects of foreclosure and find that foreclosed properties in associations have very low (3%) own-price and zero external price effects.

[9]Examples of these studies include Immergluck and Smith [19], Schuetz et al. [25] , Harding et al. [18], Campbell et al. [4], Zhang and Leonard [29]. Some of these studies weight foreclosures inversely according to their proximity to the unit for sale up to a maximum distance rather than count all proximate foreclosures equally.

[10]The delinquency measure is constructed with great difficulty because the state of delinquency is only observed on a local neighborhood basis for properties with mortgages securitized or held by Fannie Mae. Unlike this paper, the number of housing units in each area, high LTV mortgages, and variation in foreclosure time or cost are not part of the analysis.

Size and duration of foreclosure costs

In the institutional discussion of foreclosure, the diversity of the procedures across jurisdictions was noted. Judicial versus statutory foreclosure, right of redemption, eviction procedures, deficiency judgments, etc., can all influence the foreclosure process. In addition, the recent crisis prompted introduction of requirements that lenders consider modifications through the Home Affordable Modification Program and Home Affordable Refinance Program. Cordell et al. [9] document the substantial changes in the cost of foreclosure delay that occurred during the housing crisis due both to changing rules and congestion in the courts. Simply put, if foreclosure delay is important for foreclosure externalities, then there is substantial variation over time and location in the time needed for property liquidation to use in testing this effect. Cordell et al. [9] estimate average liquidation time in judicial (statutory) states at 21 (16) months at the start of the crisis in 2007 and 35 (25) months in 2013. Dagher and Sun [12] have identified a difference in mortgage terms across judicial and statutory states that presumably reflects these differential costs. Cordell and Lambie-Hanson [8] estimate the difference in foreclosure costs across states and time periods that is used as an index of foreclosure delay in this paper.

3 Additive Independence and Omitted Variables

First consider alternatives to the assumption of additive independence. To simplify the presentation, ignore heterogeneity of the physical characteristics of housing so that variation in sales price is due to time, and location, including the effects of neighborhood amenity variables. In the literature on foreclosure and sales prices, the standard specification is shown in Equation (1).

$$V_{ijkt} = \alpha + \beta F_{ijk,t-1} + \theta_{jk} + \psi_{kt} + \epsilon_{ijkt} \tag{1}$$

Here, V_{ijkt} is the logarithm of the sales price of unit i, in neighborhood (census tract or smaller) j, jurisdiction k, at time t, F is the number of foreclosures proximate to i in the time period, $t-1$, (i.e. just before t), θ is any error specific to jk, ψ is a time varying error at the jurisdiction level, and ϵ is an error uniquely associated with unit i. In this specification, θ includes the effects of omitted variables at the neighborhood and jurisdiction level and ψ includes time varying effects at the jurisdiction level. A sub-neighborhood is associated with each unit and is defined by a small circular area around the unit. There is a distinct sub-neighborhood around each unit. By saturating the specification in Equation (1) with a full set of dummy variables specific to each jk and kt combination and recalling that k is a spatial combination of the j's, any correlation between the regression error and the lagged foreclosure measure is thought to be eliminated, i.e. $E(\epsilon_{ijkt}|F_{ijk,t-1}) = 0$. The regression estimate of β is then well identified, i.e. potential bias due to selection into treatment is eliminated. Essentially, specification with a full set of jk, kt, dummies means that the relation between foreclosure and value is based on differences in the numbers of foreclosures in the individual sub-neighborhoods within j surrounding two proximate units at a given time. Alternatively it is sufficient to argue that estimates of β are unbiased because, if another variable that could cause local house price effects, say L_{ijkt}, that varies across j at time t, and hence is included in ϵ_{ijkt}, it is uncorrelated with $F_{ijk,t-1}$.

The first problem with application of the implicit markets model is the assumption of additive independence. Consider the possibility that the relation between foreclosures and property values depends on the fraction of nearby units foreclosed. An alternative specification of Equation (1) is:

$$V_{ijkt} = \alpha + \beta(F_{ijk,t-1}/H_{ijk,t-1}) + \theta_{jk} + \psi_{kt} + \epsilon_{ijkt} \tag{2}$$

where H_{ijkt} is the number of housing units in sub-neighborhood i at time $t-1$. Given that H_{ijkt} is relatively constant over time, if Equation (2) were additive independent in F, omission

of H would not bias estimates of β although it is likely that foreclosures are increasing in housing so that omitted variable bias could occur. Because H_{ijkt} can be constructed from the same data sources that provide information on F_{ijkt}, failure to include it in previous research on foreclosure externalities is curious indeed. Some papers add a measure of housing density at the j or k level of spatial disaggregation as an additive independent component of the hedonic equation because density may influence or be influenced by house value.[11] Failure to include such an obvious variable is evidence that the importance of the assumption of additive independence is not widely recognized. Once again foreclosure externalities are a specific example of a general problem. For example, literature relating number of local crimes to property value, the additive independence problem involves the failure to deflate local crimes by the number of potential local victims because the probability of being victimized as opposed to the number of individuals victimized should determine house price effects.

All of the hypotheses regarding the possible relation between foreclosure and house value presented in the literature imply that the ratio of foreclosures to housing stock rather than a simple count of foreclosures has an effect on values. The supply hypothesis states that foreclosures add to sub-neighborhood supply. The importance of addition of a single unit to supply obviously depends on the size of the market. The tenure effect of a foreclosure also depends on the number of units in the sub-neighborhood. The "appraisal" hypothesis argues that appraisers are forced to use foreclosures as comparables. Clearly the importance of adding a foreclosure sale to either the supply of units listed or comparables used to appraise depends on the number of alternative units in the area. The relation between the physical externality of an additional poorly maintained unit and the investment effect should fall as the number of well-maintained units in the sub-neighborhood rises. Finally, any measure

[11]The relation between density and sales price is complex. On the demand side, lower housing density may be the source of positive neighborhood open space externalities. Alternatively, higher density can be the result of higher land values and raise the cost of housing from the supply side. In this particular data, the estimated coefficient of number of housing units inserted as an additive independent variable is not statistically significant because of the saturated locational dummy variables. The fact housing density has no additive independent effect on value in these estimates, has no bearing on its importance for modeling the relation between foreclosure and sales price.

of the external effect associated with a given fall in value would need to be scaled by the number of units in the sub-neighborhood. Accordingly, the theoretical argument for Equation (2) rather than (1) as the appropriate specification to test any of the hypotheses relating foreclosure and nearby property values, is extremely strong.

If (2) is the proper specification and H_{ij} cannot be observed, then the appropriate dummy variable specification of (2) would be:

$$V_{ijkt} = \alpha + \sum_{ij} \beta_{ij} F_{ijk,t-1} D_{ijk} + \sum_{jk} \varphi_{jk} D_{jk} + \sum_{kt} \pi_{kt} D_{kt} + \epsilon_{ijkt} \tag{3}$$

But estimates of (3) would yield an array of β's with individual i's nested within neighborhood j's. Put another way, when the relation between the variable of interest, F, and the dependent variable is not additive independent, hypothesis testing through successive differencing becomes problematic. Furthermore, investigating the likely hypothesis that Equation (2) is not even additive independent in (F/H) cannot even be contemplated unless both F and H are observed and entered into the hypothesis testing directly.[12] Empirical results presented later also demonstrate that the relation between (F/H) and property value is convex. Proponents of additive independence may argue that they only seek an estimator of average treatment effect. The problem with this position will be apparent later in this paper.

The variation over time and jurisdiction in foreclosure cost has also been omitted from previous studies. Equation (2) assumes that β is constant while the investment effect suggests that β varies inversely with the difficulty of foreclosure. If restrictions on creditors' remedies in jurisdiction k at time t can be quantified and measured by a variable, R_{kt} , the preferred specification of the test for foreclosure externalities becomes:

$$\begin{aligned} V_{ijkt} = \alpha + \beta_1 (F_{ijk,t-1}/H_{ijk}) + \beta_2 (F_{ijk,t-1}/H_{ijk}) R_{kt} \\ + \sum_{jk} \varphi_{jk} D_{jk} + \sum_{kt} \pi_{kt} D_{kt} + \epsilon_{ijkt} \end{aligned} \tag{4}$$

[12]Some papers test for and find non-linearity in the relation between foreclosures and house prices.

This is the general specification of the foreclosure externality equation that will be used here. The size of the pure foreclosure externality is based on estimates of β_1 and the added effect due to elevated foreclosure cost is reflected in β_2. Given that R is an estimate of foreclosure delay and cost, it is anticipated that estimates of β_2 will suffer from attenuation bias.

If Equation (4) is the appropriate specification for testing foreclosure externalities and Equation (1) has been used in the literature, the relation between estimates of β in (1) and β_1 in (4) is problematic. Negative and significant estimates of β in (1) in the literature have no necessary implications for the sign or significance of β_1 in (4), particularly if R can be measured precisely. Even if $\beta_2 = 0$, the estimates of β using the specification in Equation (1) only reflect an average treatment effect of increases in foreclosures. However, estimates of this average treatment effect will vary with the density of housing units being examined as well as the relation between foreclosures and housing density. In the foreclosure data used here, the foreclosure ratio is a function of the density of housing units.[13] In sum, previous estimates of a negative and significant coefficient of β in Equation (1) may not even imply that there is an average treatment effect of foreclosure if β_1 is not significant.

Omitted variable bias is inevitable in house value equations because of the extreme diversity of the housing stock. This is particularly true because sales price is based on observation of a transaction in which the condition of the housing unit itself is endogenous as sellers may conduct substantial improvements before listing the unit. Another problem is the possibility that valuation of the unit is based on the land rather than structure. Clapp and Bardos [5] and Clapp et al. [6] have established that the existence of sales where the object of the buyer is rebuilding on the site can significantly bias estimation results, particularly the estimated coefficient of structure age. Thus, even with a full set of housing unit characteristics, there is significant potential for unobserved heterogeneity among units because the site value in an alternative use is not observed.

[13] An added complication is that when squared values of foreclosure ratio are added to Equation (4), the true relation appears to be convex.

Nearby foreclosure sales are not a physical characteristic of the unit being sold. Their influence on sales price is understood based on the classic implicit markets paper by Rosen [24]. In the past, this approach has been used to evaluate the implicit valuation of public services (particularly schools), pollution, crime, restaurant quality, and visual amenity. These factors can all generate neighborhood externalities in that they are not under the control of the homeowner although they may influence the owner's decision to maintain and sell the unit. As a practical matter, many of these neighborhood factors are excluded from empirical work that applies Rosen's implicit markets model because the burden of measuring the quality and proximity of all local amenities so that they could be included explicitly in the analysis would be overwhelming. Instead research, and certainly research on the relation between foreclosure activity and house prices, relies on differences over time and space to remove the effects of omitted neighborhood amenity variables.

The literature on foreclosure and house prices has taken the view that differencing across time and space, dividing the city into very small sub-neighborhoods, eliminates the influence of omitted variables because there is no omitted variable highly correlated across space and time with foreclosure. One exception to this is recent work by Gerardi et al. [16] who argue that delinquency in mortgage payments is highly correlated with foreclosure because delinquency is necessary for the foreclosure process to begin. They assume additive independent effects of foreclosures and serious delinquency and find that both variables have statistically significant negative effects on sales prices of nearby housing.

The general point about micro variation of omitted variables made here is that any attempt to argue that the implicit markets model may be applied should include a careful examination for omitted variables whose variation over space and time is not eliminated by differencing. Two variables are used to illustrate the micro variation problem here. One is properties with high current loan to value ratio (CLTV) because foreclosures are a subset of this variable and the incentive to maintain a unit is diminished by high CLTV giving rise to a potential external effect on house prices. The other is properties with serious delinquencies

in mortgage payments, for example properties where a foreclosure certification has been sent. Foreclosure certification is indicative of serious delinquency and is itself a first step to foreclosure which should have the same variation over time and space.[14]

There are strong theoretical reasons to believe that the number of units with high CLTV or seriously delinquency influences sales prices because, like foreclosure, high CLTV and delinquency should have a negative effect on maintenance of housing. If the relation between foreclosure and value is based on a localized physical externality due to appearance, high CLTV and delinquency should influence sales price through a similar channel. This implies that Equation (4) be written as:

$$\begin{aligned} V_{ijkt} = \alpha + \beta_1(F_{ijk,t-1}/H_{ijk}) + \beta_2(F_{ijk,t-1}/H_{ijk})R_{kt} \\ + \lambda(L_{ijkt}/H_{ijk}) + \sum_{jk} \varphi'_{jk} D_{jk} + \sum_{kt} \pi'_{kt} D_{kt} + \epsilon'_{ijkt} \end{aligned} \tag{5}$$

Just as the count of foreclosure sales around each property i is unique to a sub-neighborhood around i, the count of high CLTV and seriously delinquent units is sub-neighborhood specific. Indeed foreclosed properties are a subset of high CLTV and delinquent units. Removing the effect of L_{ijkt}, which will be used to represent either high CLTV or delinquent units, from the list of omitted variables modifies φ, π, and ϵ as indicated by the primes above these variables in Equation (5). The investment effect suggests that $\beta_1, \lambda < 0$.[15] Estimating Equation (4) with high CLTV or delinquent units as omitted variables results in $\epsilon_{ijkt} = \lambda(L_{ijkt}/H_{ijk}) + \epsilon'_{ijkt}$ and $E(\epsilon_{ijkt}|(L_{ijkt}/H_{ijk}) < 0)$. If, as seems likely, L_{ijkt} and $F_{ijk,t-1}$ are positively correlated, this would create a classic omitted variable bias problem. The resulting downward bias in estimates of β_1 survives even in a model saturated with dummy variables. The resulting bias overstates the size of the externality. More more central to the question of differences deceiving raised in this paper, omission of these variables changes the interpretation of the

[14] As noted in Gerardi et al. [16] spatial measures of mortgage delinquency are extremely difficult to construct.

[15] This omitted variable issue also applies to estimation of β in Equations (1) and (2).

effects of foreclosure on property values. Foreclosure eventually eliminates high CLTV and seriously delinquent properties as the units are resold to owners with the incentive and ability to maintain them properly.

Even if serious delinquency and high CLTV are not sufficiently correlated with foreclosure to bias the estimate of β_1 in Equation (5) or β in Equations (1) or (2), use of a specification which ignores the role of these variables creates serious problems of model interpretation as will apparent when the estimation results are discussed. Furthermore the need to consider bias or problems of interpretation due to this type of micro variation can be extended to other applications of the implicit market model. For example, crime could be substituted for foreclosure and the problem of micro variation would involve local police effort rather than high CLTV or delinquency. In the literature on pollution and environmental quality the micro variation in maintenance of surrounding property in the presence in of differentially attractive environmental conditions could be updated. Clearly the problem of micro variation in important omitted variables could apply to a vast literature relating various treatments to property values.

4 Data Preparation and Description

There is nothing remarkable about the data sources used in this study. The objective is to demonstrate that standard datasets from previous studies could have been used to conduct the analysis presented here if there had been a concern for problems of additive independence and micro variation. The variables added to the standard specification of Equation (1) were all collected from or constructed using readily available data sources.

The main data source is Corelogic Data. Corelogic (previously known as "DataQuick") is one of several vendors maintaining a large database of property transactions covering a substantial portion of the United States. Two data files, the taxation and foreclosure files, were provided by Corelogic for five parts of the Washington, D.C. metropolitan area. These

areas include the District of Columbia proper, and the counties of Montgomery and Prince George in Maryland, plus Arlington and Fairfax counties in Virginia. The taxation data file includes the most recent tax records (2015 and 2016) along with the most recent two transactions in the past for all residential and commercial properties in the area.

Following the majority of previous studies, this research focuses solely on single family units. Arms-length sales can be traced back to 2006, but, due to limitations of the foreclosure data explained below, only sales after 2008 are included. Property characteristics (living area, baths, unit age, etc.) were extracted from the most recent tax records. Approximately 1% of sales records were excluded due to missing characteristics. Sales records were also excluded if sales prices were below the 1st or above the 99th percentile in the overall sample. After data de-duplication and cleaning, there were 114,369 single family non-distressed sales ranging from 2008 to 2015 in the final estimation sample. All sales records were geocoded by parcel address for counting nearby units. Summary statistics of these sales prices and housing characteristics are reported in the top panel of Table 1. Insofar as possible, these data were treated in a fashion similar to previous foreclosure externality studies.

The second data file, foreclosure data, provides all historical property-level foreclosure records from initial notice of default to final disposition. Foreclosure units in this study refer to title transfer sales following initial notice of default. Therefore, foreclosure units in this study mostly reflect completed foreclosure proceedings, which is consistent with most of the literature.[16] The foreclosure disposition date ranges from 2006 to 2015. The common practice in the literature is to count foreclosure units within specified time and distance intervals from recent non-distressed sales. Foreclosures from the past 12 months are found to have a significant negative effect on non-distressed sales. Recent studies (Gerardi et al. [16] and Annenberg and Kung [1]) show that REO sales within one half mile have significant impact on sales price. Based on these previous findings, foreclosures from both the past year and the year before that are identified for each non-distressed sale. In addition, total nearby

[16] Approximately 94% of these foreclosures are disposed by lenders through REO sales.

units within one half mile radius are also computed in order to test the additive independence hypothesis laid out in previous section. The total surrounding unit computation is based on the most recent tax records. As shown in Table 1, total surrounding units within a half-mile radius vary from 6 to over 4000. Regressing surrounding foreclosed units in the past year against total units produces a slope coefficient of 0.008 which is significant at 1% level.[17] Further examination reveals that this relation is in fact highly nonlinear.[18]

Units with high CLTV could have a negative effect on non-distressed sales similar to foreclosed units. High CLTV is characterized by micro variation in that it is distributed over space and time in a manner similar to foreclosure. But CLTV is not observed for unsold units. An estimate of the number of high CLTV units surrounding each observed sale was constructed using the recently available FHFA zip-level annual repeat sales housing index.[19] Summary statistics relating characteristics of surrounding units as fractions of total units are also reported in Table 1. On average, foreclosure units account for less than 1% of all nearby units. However this fraction ranges as high as 10% and it has a mass point at zero.

5 Empirical Strategy and Estimation Results

Testing for additive independence

The first step in the analysis is to estimate a foreclosure externality model following procedures common in the literature. This involves estimating Equation (1) with sub-neighborhoods set at the census tract level and foreclosures within 0.5 miles of a non-distressed property sale assumed to have an additive independent relation to the logarithm of sales price. The estimated coefficient of foreclosure count in Table 2 is -0.0037 and it is statistically signifi-

[17] This result is obtained by including census tract FE. Following the custom in the literature, standard errors are clustered at the census tract level. The positive relation however does not depend on inclusion of census tract FE.

[18] Additional analysis can be found in the online appendix of this paper.

[19] Details of this housing index can be found in Bogin et al. [3]). The approach to measuring high CLTV is similar to Biswas and Davidoff [2]. Detailed discussion of calculation is included in the online appendix.

cant. At the mean of the sample, 8 foreclosures per unit sold, this implies that foreclosure externalities were lowering prices by 3%. Foreclosure rates were very high during this period and while the effect of an additional foreclosure on sales price is lower than most other studies, the effect at the mean of the sample is typical.

The first modification of Equation (1) relaxes the assumption of additive independence by replacing foreclosure count with the ratio of foreclosures to housing units in the nearby area. The estimated coefficient in Table 2 column 2 is best compared to the coefficient of foreclosure count by noting the effect at the mean of the sample, which is 0.64 multiplied by the estimated coefficient of -6.16 for foreclosure ratio. The implied reduction in property value at the mean of the sample is 3.94%. Most importantly, when an encompassing test is performed with both the number and ratio of foreclosures to housing units are both forced into the same equation, column 3 of Table 2, the estimated coefficient of foreclosure count is highly non-significant while foreclosure ratio retains its magnitude and significance.[20] As suspected, the local fraction of foreclosed properties rather than the number is related to nearby sales prices and the effect of foreclosure is not additive independent as assumed throughout the previous literature.[21]

The results in column 4 of Table 2 illustrate the critical nature of the assumption of additive independence. In this specification the number of housing units is inserted but the specification is additive independent. The results suggest that considering the number of units has no significant effect on the measured foreclosure externality. Indeed, the estimated coefficient of number of foreclosures shows a slightly, but not significantly, larger negative effect and number of units has a positive coefficient perhaps reflecting the effects of higher land values on density. Comparing the results in columns 3 and 4, the deceptive effects of the

[20] Vuong's test (Vuong [27]) for non-nested model selection also rejects the null hypothesis that models in columns 1 and 2 in Table 2 are equivalent. In the case of linear models, this likelihood ratio test comes down to a comparison between squared residuals of the two alternative models. This test is in favor of the alternative hypothesis that model in column 2 is preferred to column 1 at a P-value equal to zero up to the third decimal point.

[21] As discussed before, the relation between foreclosure and housing units is nonlinear. The fourth column of Table 2 demonstrates that entering foreclosure and housing units as additive independent variables does not solve the specification problem.

additive independence assumption are clear. But this is only the beginning of the problem.

Cordell and Lambie-Hanson [8] report substantial variation in foreclosure cost over time and across jurisdictions as the housing crisis deepened. Foreclosure cost increases with foreclosure time. As the institutions dealing with the foreclosure process became increasingly congested and the regulatory environment was modified to make foreclosure more difficult, foreclosure costs rose even in states allowing statutory foreclosure. If there is an investment effect of foreclosure, it should vary directly with foreclosure time and cost. Past studies of foreclosure externalities have generally only allowed disaggregation into judicial versus statutory foreclosure states. A measure of foreclosure cost is taken from Appendix A in Cordell and Lambie-Hanson [8]. These foreclosure costs are plotted in Figure 1. Foreclosure costs vary across the three states in this study, although all three are generally classified as statutory foreclosure states.[22]

The effect of foreclosure time and cost is tested by estimating Equation (4). Results of this exercise are shown in Table 3 where the preferred specification based on the share of foreclosed units is shown in column 2 of the table. Compared to column 1 introduction of foreclosure cost has caused the pure foreclosure effect to fall substantially. Breaking the foreclosure effect into a pure foreclosure effect, β_1, and one proportional to cost, β_2, substantially lowers estimates of the pure foreclosure effect. This implies that most of what has been previously called a foreclosure externality is due to the time that units spend in the foreclosure process rather than the behavior of creditors or some other aspect inherent in the process itself. On the other hand, results in columns 3 and 4, where foreclosure counts instead of shares are interacted with foreclosure cost, lead to totally different conclusions. The pure foreclosure effect remains almost unchanged when the interaction term is included

[22]Generally Maryland, Virginia, and the District of Columbia are classified as non-judicial foreclosure states. In Virginia and the District of Columbia this implies that the foreclose procedures can be the responsibility of the trustee if so designated in the trust. However, in Maryland, two methods of foreclosure are possible, power of sale foreclosure in which the trustee sells the property and assent to decree foreclosure where the borrower agrees to allow the court to sell the property. In either case, Maryland requires that the lender or trustee file a lawsuit in court to foreclose. The presence of either of these provisions in the mortgage greatly expedites the judicial process. Thus Maryland is usually classified as a statutory foreclosure state but it is occasionally listed as judicial.

in the model. The results in columns 3 and 4 give the false impression that foreclosure cost does not contribute significantly to the effect of foreclosure on housing price. Clearly the assumption of additive independence distorted estimates of the pure foreclosure effect and the critical role of foreclosure cost.

Testing potential omitted variables

The first two columns of Table 4 illustrate the effects of adding the ratio of the number of units with high CLTV to total housing units as an example of an omitted variable whose micro variation is similar to the foreclosure ratio. Addition of the ratio of high CLTV properties appears to have little effect on the estimated effect of the foreclosure ratio.[23]

In columns 3 and 4, squared terms are inserted in the model. In the units specification in column 3, although the foreclosure effect is shown to be convex indicating attenuation of the foreclosure effect, the estimated coefficient on the squared term does not have any meaningful impact on the marginal effect. In column 4, when foreclosure and high CLTV are measured as fractions of total units, the convex foreclosure effects are both economically and statistically significant. This is consistent with the investment rather than the supply hypothesis as the basis for foreclosure externalities. The marginal effect of foreclosure evaluated at the sample mean is -6.70, and the marginal effect of high CLTV is -0.48. Both of these marginal effects represent increases from the effects at the mean of the sample shown in column 2.

The theoretical argument for adding the share of nearby delinquent units to the specification is bolstered by the finding in Gerardi et al. [16], in a model with differencing identical to this paper, that the number of nearby delinquent units had an additive independent effect on sales prices and lowered the estimated foreclosure effect.[24] Table 5 shows that this is the

[23]The estimated coefficient of foreclosure is now interpreted as the incremental effect of nearby foreclosed units above that of nearby high CLTV units. Given that foreclosed units are assumed to have high CLTV, the total effect of nearby foreclosed units compared to having no foreclosed or high CLTV units nearby is based on the sum of the foreclosure and high CLTV effects. Robustness tests reported in an appendix to this report show that not counting recent foreclosures in the high CLTV units or including two years of lagged foreclosures leaves the conclusions of this paper unchanged.

[24]Gerardi et al. [16] report that the effects of their measure of nearby serious delinquency are larger than

case in the data used here. Comparing the results in columns 1 and 2 of the table, it is evident that adding the share of nearby delinquent units lowers both nearby housing sales prices and the effect of foreclosure on that value. Column 3 shows estimates of Equation (5) including the effects of adding both micro variation variables, shares of high CLTV units and delinquent units, to the preferred specification with the foreclosure share variables. In this specification all estimates agree well with prior expectations and the pure foreclosure effect, indicated by estimates of β_1, is now non-significant. In spite of previous evidence to the contrary, estimates of the pure average treatment effect of foreclosure have been reduced to zero. This illustrates the point made above, that a negative and significant estimate of β in Equation (1), has no necessary implication for the significance of estimates of β_1 in Equations (4) or (5).

One other specification issue that has not been considered previously is the possibility that some non-distressed sales might be short sales, which typically occur at a lower price. To the extent that short sales are spatially correlated with foreclosures, this could be a factor in producing a negative association between sales price and any measure of nearby property distress. Short sales are not directly observed but previous research has identified these as cases where the sales price be less than the original mortgage amount.[25] Comparing columns 3 and 4 of Table 5 shows that adding a short sale dummy variable to the specification of Equation (5) has no significant effect on the other parameter estimates of the model. The implied 14-15% discount for short sales is sizable but smaller than the usual estimates of the own price effect of a foreclosure sale.

Table 6 is divided into two panels to show the dramatic effect of the assumption of additive independence on estimates of treatment and/or average treatment effects. The first three columns of the upper panel show estimation results from increasingly complete specifications of Equation (5). The estimates of β_1 are initially large and significant in column

the effects of REO sales on property values. Unfortunately they do not test the hypothesis that adding delinquency lowers the estimated effect of REO sales on nearby property values.

[25] Alternatively, besides the LTV condition, the time of sale is also restricted to be less than 3 years from loan origination. The results remain very similar with this alternative definition of short sale indicator.

1 but are highly non-significant by column 3 indicating that a properly specified foreclosure externality equation produces an average treatment effect of zero. The lower panel of Table 6 has estimates based on the assumption of additive independence. Foreclosure cost cannot be a separate argument of the equation because it is a linear combination of the jurisdiction and time dummies. While the size and significance of β, the estimated coefficient of the foreclosure count, is reduced as variables are added to the specification, in the end it is still negative and significant. This illustrates the influence of the additive independence assumption for testing hypotheses about the size and existence of treatment effects. Forcing all of the important variables into the estimating equation not only produces biased estimates of the relation between additional foreclosures and property values, it allows type 1 errors, i.e. false positive indications of significant average treatment effects.

6 Conclusions and Implications

This research demonstrates the importance of the assumption of additive independence and the omission of potential important housing finance variables in testing for the relation between foreclosure and local property values. More generally, it illustrates that problems related to failure of additive independence and micro variation should be of at least as much concern as attempts to avoid bias due to selection into treatment in studies using differences in differences, propensity scores, or synthetic controls to overcome identification problems.

Clearly, the appropriate specification for the relation between foreclosure activity and nearby housing price is not additive independent. It is based on the ratio of foreclosures to housing units and allowance for the time to foreclosure must be made. When modeled as a fraction of foreclosures the convexity of the effect of foreclosure ratio on property price is apparent. The role of foreclosure delay in creating the foreclosure externality effect is substantial. This is most important because one reaction to the publication of studies showing large foreclosure externalities has been regulations that raise the cost of foreclosure, a classic

case of a regulatory reaction that makes the initial problem worse.

The empirical results also confirm that the fraction of high CLTV properties has a negative relation to nearby sale prices but this does not materially influence the estimated coefficient of foreclosure ratio. Thus the traditional concern that selection into foreclosure is based on high CLTV which has a negative relation to property value does not seem warranted. However, there are major implications of the finding that high CLTV also depresses property values because foreclosure, by recapitalizing the housing unit, reduces the stock of high CLTV properties. Thus foreclosure is a treatment for the effect of high CLTV on nearby housing and without that treatment the stock of such housing would tend to increase. Accordingly, foreclosure may have a temporary negative local house price effect but it is the cure for a potentially long term problem of high CLTV properties. This demonstrates that the total derivative with respect to a treatment may differ from the partial derivative or that it is easy to confuse a treatment with a disease, particularly if, in the short run, the treatment makes the patient weaker.

Addition of the ratio of seriously delinquent properties which also micro vary with foreclosure, leaves estimates of the pure foreclosure effect non-significant. This is the classic problem of an omitted variable that is correlated with both the treatment and the outcome. Taken together, the effects of the high CLTV and seriously delinquent ratio variables demonstrate the importance of understanding the full empirical relation between treatment and outcomes.

The possibility of using an instrument for foreclosure has not been discussed here. Use of an IV estimator might help if the problem was selection into treatment. But that is not the concern here. The instrument would not deal with the problems arising from the assumption of additive independence and finding an instrument not correlated with the omitted variables, at least in the case of high CLTV and serious delinquency, does not appear possible. Furthermore the problem with omitting these variables is that our understanding of the entire process relating nearby foreclosures and property values is fundamentally changed

when the omitted variables are included in the model. IV estimates that exclude these variables would also be deceptive.[26]

A more general lesson to be drawn from this exercise is that invoking differences in differences to quantify the effect of a treatment, even when there is no selection problem, requires strong assumptions. Most obvious is additive independence of the treatment effect, or at least caution in interpreting the estimated coefficient of treatment as a number that is applicable to any specific circumstance. The results also demonstrate that ignoring additive independence and claiming to measure an average treatment effect can be deceiving because the estimating equation is seriously misspecified. Less obvious is the possibility that a more fully specified model can reveal something about the mechanism underlying the treatment effect that is material to understanding the overall relation between the treatment and social outcomes. In the case of this study, all of the variables added to the model were easily available in data used by most previous studies. The failure to consider the specification used here was based on the conviction that differences in differences do not deceive. At least in the case of the foreclosure externality literature, that conviction was misplaced.

[26]Footnote 2 considers the possibility that, if $V = v(F, D, L)$ where F, D, and L are measures of nearby foreclosures, serious default, and high CLTV properties respectively, estimates of v_F may not reveal the full treatment effect of foreclosure. The empirical results presented here suggest that in the expression $\frac{dV}{dF} = v_F + v_D \frac{dD}{dF} + v_L \frac{dL}{dF}$, $v_F \approx 0$, while v_D,$v_L < 0$ which, given $\frac{dD}{dF}$, $\frac{dL}{dF} < 0$, implies $\frac{dV}{dF} > 0$.

References

[1] Elliot Annenberg and Edward Kung. Estimates of the size and source of price declines due to nearby foreclosures. *American Economic Review*, 104(8):2527–2551, 2014.

[2] Arnab Biswas and Thomas Davidoff. How does mortgage distress affect neighboring properties' resale value? evidence from worcester, ma. *Working paper presented at the Midyear Meetings of the American Real Estate and Urban Economics Association*, 2016.

[3] Alexander N. Bogin, William M. Doerner, and William Larson. Local home price dynamics: new indices and stylized facts. *FHFA Staff Working Paper Series*, 16(1), 2016.

[4] John Y. Campbell, Stefano Giglio, and Parag Pathak. Forced sales and house prices. *American Economic Review*, 101(5):2108 – 2131, 2011.

[5] John M. Clapp and Katsiaryna Salavei Bardos. Hedonic pricing with redevelopment options: a new approach to estimating depreciation effects. *Journal Of Urban Economics*, 67(3):362–377, 2010.

[6] John M. Clapp, Katsiaryna Salavei Bardos, and S.K. Wong. Empirical estimation of the option premium for residential development. *Regional Science and Urban Economics*, 42(1-2):240–256, 2012.

[7] Terrence M. Clauretie and Nasser Daneshvary. Estimating the house foreclosure discount corrected for spatial price interdependence and endogeneity of marketing time. *Real Estate Economics*, 37(1):43 – 67, 2009.

[8] Larry Cordell and Lauren Lambie-Hanson. A cost-benefit analysis of judicial foreclosure delay and a preliminary look at new mortgage servicing rules. *Journal of Economics and Business*, 84:30 – 49, 2016.

[9] Larry Cordell, Liang Geng, Laurie S. Goodman, and Lidan Yang. The cost of foreclosure delay. *Real Estate Economics*, 43(4):916–956, 2015.

[10] N. Edward Coulson and Herman Li. Measuring the external benefits of homeownership. *Journal of Urban Economics*, 77:57 – 67, 2013.

[11] Lin Cui and Randall Walsh. Foreclosure, vacancy, and crime. *Journal of Urban Economics*, 87(2):72–84, 2015.

[12] Jihad Dagher and Yangfan Sun. Borrower protection and the supply of credit: Evidence from foreclosure laws. *Journal of Financial Economics*, 121(1):195 – 209, 2016.

[13] Angus Deaton and Nancy Cartwright. Understanding and misunderstanding randomized controlled trials. *NBER Working Paper*, (22595), 2016.

[14] Lynn M. Fisher, Lauren Lambie-Hanson, and Paul Willen. The role of proximity in foreclosure externalities: Evidence from condominiums. *American Economics Journal: Economic Policy*, 7(1):119 – 140, 2015.

[15] W. Scott Frame. Estimating the effects of mortgage foreclosures on nearby property values: a critical review of the literature. *Federal Reserve of Atlanta Economic Review*, November:1 – 9, 2010.

[16] Kristopher Gerardi, Eric Rosenblatt, Paul S. Willen, and Vincent Yao. Foreclosure externalities: new evidence. *Journal of Urban Economics*, 87:42 – 56, 2015.

[17] Jeremy R. Groves and William H. Rogers. Effectiveness of rca institutions to limit local externalities: using foreclosure data to test covenant effectiveness? *Land Economics*, 87(4):559 – 581, 2011.

[18] John P. Harding, Eric Rosenblatt, and Vincent W. Yao. The contagion effect of foreclosed properties. *Journal of Urban Economics*, 66:164 – 178, 2009.

[19] Dan Immergluck and Geoff Smith. The external costs of foreclosure: The impact of single-family mortgage foreclosures on property values. *Housing Policy Debate*, 17(1): 57 – 70, 2006.

[20] Lauren Lambie-Hanson. When does delinquency result in neglect? mortgage distress and property maintenance. *Journal of Urban Economics*, 90:1 – 16, 2015.

[21] Atif Mian, Amir Sufi, and Francesco Trebbi. Foreclosures, house prices, and the real economy. *The Journal of Finance*, 70(6):2587–2612, 2015.

[22] Anthony Pennington-Cross. The value of foreclosed property. *Journal of Real Estate Research*, 28(2):193 – 214, 2006.

[23] Kristin L. Perkins, Michael J. Lear, and Elyzabeth Gaumer. Concentrated foreclosure activity and distressed properties in new york city. *Urban Affairs Review*, pages 1–30, 2016.

[24] Sherwin Rosen. Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, 82(1):34–55, 1974.

[25] Jenny Schuetz, Breen Vicki, and Ingrid Gould. Neighborhood effects of concentrated mortgage foreclosures. *Journal of Housing Economics*, 17(4):306 – 319, 2008.

[26] Albert Sumell. The determinants of foreclosed property values: Evidence from inner-city cleveland. *Journal of Housing Economics*, 18(2):45 – 61, 2009.

[27] Quang H. Vuong. Likelihood ratio tests for model selection and non-nested hypothesis. *Econometrica*, 57(2):307–333, 1989.

[28] Stephan Whitaker and Thomas J. Fitzpatrick. Deconstructing distressed-property spillovers: The effects of vacant, tax delinquent, and foreclosed properties in housing submarkets. *Journal of Housing Economics*, 22(2):79 – 91, 2013.

[29] Lei Zhang and Tammy Leonard. Neighborhood impact of foreclosure: a quantile regression approach. *Regional Science and Urban Economics*, 48:133 – 143, 2014.

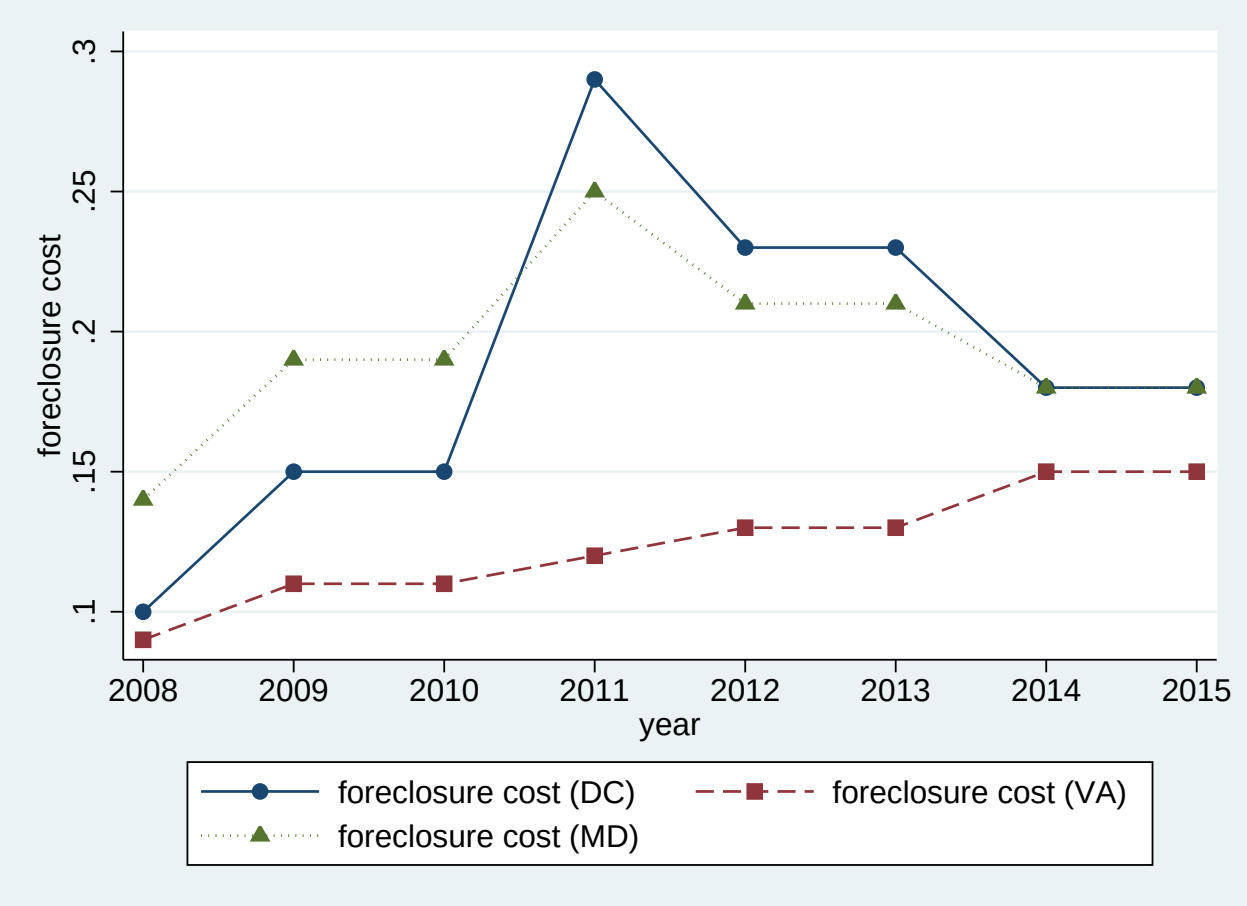

Figure 1: Foreclosure Cost by State and Year

Note: Foreclosure costs are taken from Appendix A in Cordell and Lambie-Hanson [8].

Table 1: Summary Statistics

	Mean	Std.dev.	Min	Max	Observations
Panel 1: Single family non-distressed sales unit information					
Single family non-distressed sales price	462399.30	265641.90	75000.00	2395000.00	114369
Unit age	43	27	1	133	114369
Lot size (square foot)	4436.72	2852.70	1117.00	9998.00	114369
Living area (square foot)	1761.57	681.05	832.00	6114.00	114369
Number of bath rooms	3.05	1.01	1	5	114369
Fireplace indicator	0.65	0.48	0	1	114369
Panel 2: unit counts within one half mile radius					
All surrounding units	1260.16	569.24	6	4215	114369
Foreclosure units in the past year	8.34	9.87	0	178	114369
Foreclosure units in the year before last year	7.03	9.27	0	173	114369
Delinquent units in the past year	11.80	18.18	0	235	114369
High CLTV units (at sale date)	119.91	97.24	0	866	114369
Fraction of foreclosure units in the past year	0.64%	0.65%	0	10.06%	114369
Fraction of foreclosure units in the year before past year	0.54%	0.62%	0	9.86%	114369
Fraction of denquent units in the past year	0.92%	1.21%	0	20.27%	114369
Fraction of high CLTV units (at sale date)	9.84%	7.29%	0	81.70%	114369

Table 2: Testing the Additive Independence Assumption: Number of Units

Dependent variable:	log of sale price			
	(1)	(2)	(3)	(4)
Number of foreclosures in the past year	-0.0037		0.0000	-0.0043
	(0.0002)***		(0.0004)	(0.0003)***
Share of foreclosures in the past year		-6.1590	-6.1143	
		(0.2946)***	(0.6440)***	
Number of total nearby units				0.00005
				(0.0000)***
Log of lot size	0.1442	0.1444	0.1444	0.1445
	(0.0038)***	(0.0038)***	(0.0038)***	(0.0038)***
Log of living area	0.3380	0.3373	0.3373	0.3380
	(0.0078)***	(0.0077)***	(0.0077)***	(0.0077)***
Number of baths	0.0484	0.0484	0.0484	0.0484
	(0.0019)***	(0.0019)***	(0.0019)***	(0.0018)***
Fireplace indicator	0.0534	0.0530	0.0530	0.0528
	(0.0030)***	(0.0030)***	(0.0030)***	(0.0030)***
Unit age	-0.0114	-0.0114	-0.0114	-0.0115
	(0.0005)***	(0.0005)***	(0.0005)***	(0.0005)***
Unit age squared	0.0001	0.0001	0.0001	0.0001
	(0.0000)***	(0.0000)***	(0.0000)***	(0.0000)***
Constant	9.2674	9.2848	9.2848	9.2135
	(0.0529)***	(0.0526)***	(0.0526)***	(0.0542)***
Sale year and jurisdiction FE	Y	Y	Y	Y
Census tract FE	Y	Y	Y	Y
R^2	0.868	0.868	0.868	0.868
Observations	114369	114369	114369	114369

Table 3: Tesing the Additive Independence Assumption: Foreclosure Cost

Dependent variable:	log of sale price			
	(1)	(2)	(3)	(4)
Number of foreclosures in the past year			-0.0032	-0.0035
			(0.0005)***	(0.0005)***
Number of foreclosures*foreclosure cost			-0.0028	-0.0050
			(0.0029)	(0.0030)*
Share of foreclosures in the past year	-6.1590	-2.3464		
	(0.2946)***	(0.6527)***		
Share of foreclosures*foreclosure cost		-24.6210		
		(4.4253)***		
Number of total nearby units				0.00005
				(7.10e-06)***
Constant	9.2848	9.2856	9.2424	9.1876
	(0.0526)***	(0.0525)***	(0.0530***	(0.0543)***
Control variables	Y	Y	Y	Y
Sale year and jurisdiction FE	Y	Y	Y	Y
Census tract FE	Y	Y	Y	Y
R-squared	0.868	0.868	0.868	0.868
Observation	114369	114369	114369	114369

Note: Same set of control variables as in Table 2 is used here.

Table 4: Testing for Micro Variation of High CLTV and Convexity

Dependent variable:	log of sale price			
	(1)	(2)	(3)	(4)
Number of foreclosures in the past year	-0.0034		-0.0042	
	(0.0003)***		(0.0004)***	
Number of foreclosures squared			0.00001	
			(0.0000)***	
Share of foreclosures in the past year		-5.8431		-6.8780
		(0.3302)***		(0.4288)***
Share of foreclosures squared				27.7809
				(5.6355)***
Nunber of high CLTV units	-0.0001		-0.0001	
	(0.0000)**		(0.0001)	
Number of high CLTV units squared			0.0000	
			(0.0000)	
Share of high CLTV units		-0.1251		-0.5764
		(0.0554)**		(0.0756)***
Share of high CLTV units squared				0.9933
				(0.1301)***
Constant	9.3481	9.3038	9.3532	9.3546
	(0.0531)***	(0.0523)***	(0.0529)***	(0.0522)***
Control variables	Y	Y	Y	Y
Sale year and jurisdiction FE	Y	Y	Y	Y
Census tract FE	Y	Y	Y	Y
R^2	0.868	0.868	0.868	0.869
Observations	114369	114369	114369	114369

Note: Same set of control variables as in Table 2 is used here.

Table 5: Testing for Micro Variation of High CLTV and Delinquency

Dependent variable:	log of sale price			
	(1)	(2)	(3)	(4)
Share of foreclosures in the past year	-2.2492	-0.8077	-0.5423	-0.5272
	(0.6449)***	(0.7337)	(0.7586)	(0.7749)
Share of foreclosures*foreclosure cost	-23.3606	-30.4990	-30.8290	-32.4830
	(4.4875)***	(4.6043)***	(4.6854)***	(4.8539)***
Share of high CLTV units			-0.1050	-0.1221
			(0.0545)***	(0.0556)**
Share of DLQ units in the past year		-0.6999	-0.5946	-0.6110
		(0.2011)***	(0.2063)***	(0.2125)***
Short sale indicator	-0.1412	-0.1411	-0.1410	
	(0.0056)***	(0.0056)***	(0.0056)***	
Constant	9.3557	9.3691	9.3830	9.3159
	(0.0528)***	(0.0529)***	(0.0527)***	(0.0522)***
Control variables	Y	Y	Y	Y
Sale year and jurisdiction FE	Y	Y	Y	Y
Census tract FE	Y	Y	Y	Y
R-squared	0.876	0.876	0.876	0.869
Observation	114369	114369	114369	114369

Note: Same set of control variables as in Table 2 is used here.

Table 6: Effects of Assuming Additive Independence

Panel A: testing without additive independence assumption			
Dependent variable:		log of sale price	
	(1)	(2)	(3)
Share of foreclosures in the past year	-5.5927	-1.6663	-0.5423
	(03097)***	(0.6820)***	(0.7586)
Share of foreclosures*foreclosure cost		-25.0576	-30.8290
		(4.3647)***	(4.6854)***
Share of high CLTV units	-0.1085	-0.1269	-0.1050
	(0.0542)**	(0.0549)**	(0.0545)***
Share of DLQ units in the past year			-0.5946
			(0.2063)***
Short sale indicator	-0.1412	-0.1410	-0.1410
	(0.0056)***	(0.0056)***	(0.0056)***
Constant	9.3714	9.3749	9.3830
	(0.0528)***	(0.0527)***	(0.0527)***
Control variables	Y	Y	Y
Sale year and jurisdiction FE	Y	Y	Y
Census tract FE	Y	Y	Y
R-squared	0.876	0.876	0.876
Observations	114369	114369	114369
Panel B: testing with additive independence assumption			
Dependent variable:		log of sale price	
	(1)	(2)	(3)
Number of foreclosures in the past year	-0.0035	-0.0026	-0.0010
	(0.0003)***	(0.0005)***	(0.0005)**
Number of foreclosures*foreclosure cost		-0.0057	-0.0143
		(0.0029)**	(0.003)***
Number of high CLTV units	-0.0002	-0.0002	-0.0001
	(0.00004)***	(0.00004)***	(0.00004)***
Number of DLQ units in the past year			-0.0009
			(0.0002)***
Number of total nearby units	0.0001	0.0001	0.0001
	(7.63e-06)***	(7.68e-06)***	(7.68e-06)***
Short sale indicator	-0.1411	-0.1410	-0.1408
	(0.0056)***	(0.0056)***	(0.0056)***
Constant	9.3583	9.3295	9.3275
	(0.0544)***	(0.0544)***	(0.0543)***
Control variables	Y	Y	Y
Sale year and jurisdiction FE	Y	Y	Y
Census tract FE	Y	Y	Y
R-squared	0.876	0.876	0.876
Observations	114369	114369	114369

Note: Same set of control variables as in Table 2 is used here.