The cost of misinformation: Evidence from an industrial disaster*

Marianne Bléhaut, CREST †

January 17, 2017

PRELIMINARY. Please do not cite.

Abstract

How accurate is the assumption of perfect information in real markets? In particular, for complex goods such as housing, quality is possibly unobserved. This would lead to an inefficient allocation of resources and a welfare loss. Using an industrial disaster as quasi-experimental setting, I measure the housing market reaction to a symmetric information shock and show by contrast that the initial situation was characterized by imperfect information. More precisely, I show that when information relative to bad environmental quality is revealed, housing prices significantly and robustly decrease in the short term. Far from being a mere transitory shock, this effect is reinforced in the medium term. All in all, I find a 1 to 2% price decrease in the short term, and an additional decrease twice as large in the medium term. In addition, the relative number of vacant housing units increases in these areas, and there is some evidence supporting residential sorting since mean household earnings decrease. The literature relative to housing reaction to environmental quality still heavily relies on hedonic methods and thus on the assumption of perfect information. This paper shows that this assumption should be treated with caution since in many settings it is likely not to be valid.

Keywords: housing market, industrial hazard, natural experiment, propensity score matching.

JEL classification: R23, R21, R20, C21.

*I am deeply indebted to Laurent Gobillon and Miren Lafourcade for their support and guidance of this work. I am also very grateful to Xavier d'Haultfoeuille for his patience and thoughtful advice. I am very thankful to Luc Behaghel, Gabrielle Fack, Florence Goffette-Nagot, Christian Hilber and Pierre Picard for their excellent comments on this work, and to Gabriel Ahlfeldt, Giuseppe Arbia, Sylvain Chabé-Ferret, Eric Gautier, Stephen Gibbons, Arthur Grimes, Stephan Heblich, Camille Hémet, Alessandro Iaria, Manasa Patnam, Roland Rathelot, Olmo Silva, Yves Zénou, Yanos Zylberberg for useful discussions and comments. I would also like to thank conference participants at the SERC Annual Conference, the Annual North American Meeting of the Regional Science Association International, the Congress of the European Regional Science Association, the Annual Meeting of the French Economic Association and World Conference of the Spatial Econometrics Association for their comments.

†Please address correspondance to marianne.blehaut@ensae.fr.

1 Introduction

How accurate is the assumption of perfect information in real markets? In practice, if individuals are not well-informed, microeconomic theory predicts an inefficient allocation of resources and a welfare loss. There is evidence of imperfect information settings in a wide array of situations. There is for example a literature pointing to imperfect information in the context of health risk perception, insurance choice or schooling decisions. Indeed, Viscusi, (1990) shows that consumers have a positive bias when assessing cancer risk related to smoking, which then reduces the probability of smoking. Duflo and Saez, (2003) highlight the role of information in individuals' choice of retirement plan. School quality information also leads to substantial changes in enrollment decisions of lower-income children, as shown by Hastings and Weinstein, (2008), while Jensen, (2010) highlights the role of *perceived* returns to schooling (as opposed to objective returns) in schooling decision in Dominican Republic. What these examples have in common is twofold. First, each case is characterized by uncertain future utility for decision-making individuals. Second, the decision is multi-dimensional, in the sense that it may affect present and future utility through many channels.

The housing market combines those two aspects, as housing choice will typically determine future utility and comprises many components (such as location, size, number of rooms, local amenities etc.). In particular, the environmental quality of a given area is likely to be partially unknown, and its perception may be biased by the presence of other amenities or nuisances. Gayer, Hamilton, and Viscusi, (2000) thus cites a 1987 report comparing expert environmental risk assessment to people's perception, which finds a discrepancy in risk ranking. More precisely, people seem to overstate the level of risk, and Gayer, Hamilton, and Viscusi find that information released with official assessment seems to decrease the perceived level of risk. In some cases on the contrary, additional information leads to a an increased perception of risk. For example, Pope, (2008) shows that mandatory seller disclosure of risk status has a negative impact on property values.

However, the literature aiming at measuring taste or distaste for environmental quality still heavily relies on Rosen's (1974) seminal paper on hedonic models. The general principle underlying hedonic model applications is that all else held equal, the taste or distaste for a given amenity is revealed through higher or lower housing prices. The price variation can then be used as a measurement of marginal willingness to pay for this amenity under a set of assumptions that include perfect information. As noted by Parmeter and Pope, (2012), assessing preferences for environmental quality based on housing market prices has gained tremendous popularity in the empirical literature and there are countless applications of this principle to measure marginal willingness to pay for a wide array of local characteristics. For example, Davis, (2004) measures the housing market reaction to a strong and unexplained rise in local cancer risk, finds a significant housing price decrease. Consistent with Rosen, (1974), he further interprets this drop as the marginal willingness to pay to avoid cancer risk of this magnitude. Similar approaches have been used to assess the housing market reaction to a variety of changes in amenities, from Superfund cleanups (Greenstone and Gallagher, 2008) to toxic plant openings and closings (Currie et al., 2015) or shale gas development (Muehlenbachs, Spiller, and Timmins, 2015).

Moreover, housing markets are not standard, and attempts at modeling and predicting housing prices often lead to the conclusion that it is not an efficient market. Case and Shiller, (1989) indeed show that there is an important time persistence of real housing prices, and that real interest rates do not seem to be incorporated in prices. These counter-intuitive results can at least partially be explained by the specificities of the housing market such as high transaction costs or tax considerations. They are also consistent with behaviors induced by the "disposition effect". Shefrin and Statman, (1985) model this tendency to "sell winners too early and ride losers too long" on financial markets even when the contrary would be more efficient, and attribute it to loss aversion. Similar behaviors have been observed on the housing market. For example, Genesove and Mayer, (2001) analyze Boston housing market in the 1990s and find evidence of nominal loss aversion among sellers. As a result of these specificities, prices alone may not be able to accurately reflect the neighborhood changes that occur when risk or risk perception shift in a given area, especially in the short term.

An additional challenge for empiricists willing to assess the impact of environmental quality is the concern of endogeneity. Indeed, the decision of building a new dangerous plant is often highly political, and such facilities tend to be constructed in neighborhoods with lower than average *ex ante* housing prices and very specific socio-demographic characteristics (Davis, 2011). This problem problem is often solved using quasi-experimental settings, in which a sudden unforeseen change modifies local amenities. The American *Superfund* clean-up program has for example been used in a number of publications (see for example Greenstone and Gallagher, 2008; Kiel and Williams, 2007; Viscusi and Hamilton, 1999). Another example is provided by legal changes in pollution control, such as the *Clean Air Acts* passed in the USA in the 1970's (see for example Chay and Greenstone, 2005; Greenstone, 2002).

This paper contributes to the existing literature in several ways. First, I provide a test of market reaction to a symmetric pure information shock. More specifically, I study the reaction of housing markets located in the vicinity of dangerous plants to the information provided by an industrial disaster. This accident was both very strong and heavily covered by local and national television and newspapers, creating a strong information shock. It did not however lead to an immediate change in policy or to a reevaluation of industrial risk by experts. In a perfect information setting, such a shock should have no impact on the housing market, thus any consequent reaction reveals individuals' reevaluation of the risk. Contrary to Pope, (2008), I argue that this shock is symmetric in the sense that both buyers and sellers can be *ex ante* oblivious to the risk status of a given property. Once environmental quality becomes observed, I show that several characteristics of the local housing markets change to reflect that differentiation and that this change is durable. This central finding is at odds with the assumption of perfect information.

Second, I rely on high-quality administrative data that allows me to analyze the housing market adjustment both in terms of transaction price and in terms of the volume of transactions and neighborhood characteristics. In particular, I am able to analyze changes in vacancy rate and average household earnings, as indicators of neighborhood composition change. Third, by analyzing medium-term effects and not only short-term adjustment, I show that the information shock has a lasting effect on housing market characteristics, and does not merely lead to transitory friction. More precisely, I define the short-term impact of the information shock as roughly one year after the disaster, and the medium-term impact as three years after the disaster.

Overall, the results of the paper are consistent with an initial imperfect information setting. I find a robust significant price decrease of about 1 to 2% in the vicinity of dangerous plants one year after the information shock. This effect is stronger in areas that are closer to the origin of the shock and that are likely to receive heavier media coverage. In addition, other characteristics of the at-risk neighborhoods vary in reaction to the change in risk perception. In particular, the vacancy rate of at-risk neighborhoods significantly increases. The average income of the households living in an at-risk area decreases, which is consistent with residential sorting. In the medium-term, the negative impact on housing prices reinforces, and there is an additional price decrease roughly twice as high as the initial one. This shows that the negative shock on risk perception is both strong and durable and generates a persistent differentiation between at-risk neighborhoods and others.

The following sections first present the empirical strategy, then the data sources and descriptive statistics. Section 4 presents the results and discussion, and section 5 concludes. The appendix gathers supplementary figures and tables.

2 Empirical strategy

2.1 A natural experiment

In September 2001, the chemical plant AZF exploded in Toulouse (south of France).[1] The plant itself was almost destroyed, and other damage for the city was both extremely strong and unexpected. Indeed, the accident amounted to the explosion of 20 to 40 tons of TNT and could be felt as far as 75km away from the site. Material damage such as broken windows occurred up to 7km from it and many amenities were destroyed. Among them, about a hundred schools and more than 26,000 housing units were damaged. More than 2,400 people were hospitalized and 31 were killed in the explosion. Figure 1 shows examples of the destruction caused by the accident to the plant itself and to other facilities. Indeed, as shown in subfigure (c), the AZF plant was almost completely destroyed. Subfigures (a) and (b) show a bus facility and a gaz station located close to the plant, and which were very severely damaged as well.

The plant was known to be highly dangerous, but the scope of the consequences was much higher than what was considered likely in accident scenarios at the time. The gap was such that at first the accident was thought to be a terrorist attack. It was soon announced that it was in fact an accident, which was widely reported in local and national press.[2] It is now considered as the worst industrial disaster in France since the Second World War. At the time of the accident, the event and its local consequences were widely covered in the media. Indeed, for a few weeks, images of the accident were heavily broadcast on television and printed in newspapers. *Le Monde*, a French national newspaper with over 400,000 total circulation in 2001, published at least three stories related to the accident every day until October, 3$^{\text{rd}}$, and at least one every day until October, 10$^{\text{th}}$. Figure 2 shows examples of images broadcasted in television news programs on September 21$^{\text{st}}$. These images show very precise examples of the material damage that followed the accident, both on the site of the plant (subfigure (b)) and further away in the city. Subfigure (a) thus shows the ringroad of the city with people walking out of badly damaged cars, while sufigures (c) and (d) show material damage in an apartment complex and an appliance store. Other images, not reported here, also showed human casualties and temporary hospitals being set up in the city.

Combined with the extreme scarcity of such events, this media coverage can reasonably lead to the assumption that industrial risk perception changed in France after September 2001. An alternative hypothesis could be that the consensus on the level of risk around dangerous plants evolved following the accident. There is however little evidence supporting that hypothesis. For example, the French Ministry of sustainable development clearly states that there was a sharp political change following the AZF accident, but notes that it mostly focused on limiting urban growth around dangerous plants and better informing the local populations, rather than reassessing risk in existing facilities or strengthening the monitoring of dangerous plants.[3] Indeed, the government institution in charge of monitoring dangerous plants (*Inspection des Installations Classées*) was created in 1976 and its missions are still similar to those of that period: monitoring the dangerous plants and enforcing the risk prevention and risk control regulation. Chabbal, (2005) also notes that ministerial circulars sent to local *Préfets* (prefects) after the AZF accident "do not introduce any radical change and even build on regular instructions towards decentralized State services".[4] In this paper, Chabbal analyzes the case of a chemical plant in Meaux (Paris area) that falls under the same regulation as the AZF plant, and was even authorized to be extended in 2002, a few

[1] The location of the plant is represented by a triangle on Figure 3.

[2] On September 25$^{\text{th}}$, the prosecutor announced in the press that the accident thesis was more likely to be true, and on October, 6$^{\text{th}}$ an independent panel of experts concluded that catastrophe was the result of an accident. This was less than a month after the disaster.

[3] Source: http://www.developpement-durable.gouv.fr/10-ans-apres-AZF-les-avancees-en,24290.html.

[4] Original text in French, translation by the author.

(a) A destroyed gaz station in front of the AZF plant

(b) A bus facility damaged by the explosion

(c) The main chimney of the AZF plant after the accident

Source: www.lemonde.fr (September 2001), ©AFP/ERIC CABANIS; NICOLAS AUER/MAXPPP; AFP/PASCAL PAVANI

Figure 1: Examples of the destruction caused by the AZF explosion

(a) Toulouse ringroad

(b) One of AZF warehouses

(c) Damaged housing

(d) An appliance store close to the AZF plant

Source: www.ina.fr, television news programs (channels France 2 and France 3 Régions Toulouse), screenshots

Figure 2: Examples of TV news program images on September 21st 2001

months after the accident. A 2014 ruling of the French Council of State[5] further confirmed that the *Inspection des Installations Classées* had indeed been monitoring the AZF plant and that the accident resulted from a combination of human mistake and rare chemical reaction that could not have been anticipated or prevented by the State.

I thus argue that it is more likely that the accident constituted a strong, unexpected national shock in risk perception, which can thus be used to identify the effect of an increase in risk perception on other at-risk areas in France.

2.2 Theoretical impact

Assuming the AZF accident changed risk perception can lead to several predictions regarding the evolution of at-risk housing markets. If households had a biased ex-ante risk perception, their behaviors will reflect their beliefs update and preferences. In the case of a negative ex-ante bias and preference for safety, the willingness to pay of potential buyers or renters is likely to decrease, causing a downward shift of the demand curve. The consequences of this shift in the short and long term then depend on the elasticity of both the demand and supply curves. In a standard case, one might expect both a decrease in prices and volume of transactions. If the housing supply is perfectly inelastic (as may be the case in the short term), the demand shift would only impact prices. If on the contrary the supply is highly elastic, the shift would mostly impact quantities.

Moreover, homeowners may be more reluctant than potential buyers to adjust their expected sale price at a lower level, as suggested by Genesove and Mayer, (2001). This would lead to a transitory discrepancy between supply and demand, and thus potentially to both a decrease of the average number of transactions and an increase of the number of vacant dwellings. The impact of the information shock could diminish with time, and the housing market could go back to its initial state. In this case, the sellers' reluctance to adjust would be rational: they would hold on to their house and wait for the shock to pass. If on the contrary the negative demand shock is persistent over time, this behavior is not rational in the sense that they would lose less by selling their property as early as possible. In such a context, there could be in the long term an increase in destruction and a reduction in construction in at-risk areas. Given the time frame of this study it is however unlikely that any effect regarding these outcomes could be measured.

The characteristics of residents in at-risk areas could also change as a consequence of the information shock. In particular, if well-off households are more mobile than modest households, increased risk perception may have a pauperization effect in the vicinity of dangerous plants. This could happen even in the hypothetical case in which preferences and risk perception are homogeneous across households. In addition, one might expect that preferences regarding risk are heterogeneous among the population. Families are often cited as a category that could be especially sensitive to environmental risk, although they also tend to face higher mobility costs, whether financial or organizational (such as finding new schools for children for example).

Finally, Davis, (2011) provides an early evidence that housing prices and rents do not react in a similar way to changes in environmental amenities, which has since been confirmed by Grainger, (2012) in a paper addressing this specific point. In particular, these papers show that rent-elasticity to environmental quality is about twice as small as price-elasticity. This result could stem from a number of theoretical channels. First, there is increasing evidence that the average renter is likely to be different from the average homeowner, in particular in terms of income, mobility and household composition (Glaeser and Gyourko, 2007). In particular, families are more likely to own a house than to rent. These features could be associated with heterogeneous preferences for environmental quality. Second, as noted by by Grainger, (2012) "unobserved expectations may play a role" since renters only pay for current environmental quality, while homeowners may capitalize

[5] December 17$^{\text{th}}$, 2014 ruling.

expected changes in environmental quality. The datasets used in this paper do not include rents, thus the predictions regarding housing market characteristics are more complex in this context. First, one can expect current homeowners to face higher mobility costs than renters, thus leading to an increase in the share of homeowners in the at-risk areas. Second, since renters can more easily move out, the vacancy rate should increase more where there are fewer homeowners. Third, the intensity of the price effect with respect to the initial share of homeowners could go both ways. Indeed, owned dwellings are likely to be as different from rented dwellings as homeowners are different from renters. In subsection 4.5, I thus study the heterogeneity of the estimated impacts with respect to the initial share of homeowners.

2.3 Identification strategy

Assessing the impact of the AZF accident can be thought of in the econometric framework developed for public policy evaluation. The Rubin model (1974) has become the standard guideline for such questions, and it is useful to recall its main features.

I denote Y_i^1 the outcome of a treated location, Y_i^0 the outcome of the same location in the absence of treatment, and T the "treatment" dummy. In our case, the treatment is the exposition to an industrial risk. The parameter of interest is the *average treatment effect on the treated*, which is the average difference between the two potential outcomes for at-risk units: $ATT = \mathbb{E}\left[Y_i^1 - Y_i^0 | T_i = 1\right]$. By definition, one cannot observe both Y_i^1 and Y_i^0 (what happened in the at-risk areas after the accident and what would have happened had the accident not occurred). This "fundamental problem of causal inference" has led to different empirical strategies in the literature.

These methods all resort to non-treated units to estimate a convincing counterfactual outcome. In the present case, this amounts to using areas that are not exposed to industrial risk in order to assess what would have happened in at-risk areas had the accident not taken place. The construction of the counterfactual is critical for the credibility of the results obtained from quasi-experimental settings. In particular, the implantation of dangerous plants is likely not to be random, thus a direct comparison of outcomes between at-risk areas and other areas would lead to biased estimates.

A natural approach would be to resort to a difference-in-difference strategy, comparing at-risk areas with control areas both before and after the accident. The ATT is then estimated as follows:

$$\widehat{ATT} = \frac{1}{n_1} \sum_{i \in I_1} \left[Y_{it'}^{T=1} - Y_{it}^{T=1}\right] - \frac{1}{n_0} \sum_{i \in I_0} \left[Y_{it'}^{T=0} - Y_{it}^{T=0}\right],$$

where t and t' denote respectively the before and after periods, I_1 and I_0 the set of at-risk (respectively control) locations, and n_1 and n_0 the number of locations in I_1 (respectively I_0). This strategy relies on the assumption that in the absence of the AZF accident, the outcome in the at-risk and control groups would have followed the same trend. Whether this "common trend assumption" holds can only be checked before the accident. It is however likely to be rejected since areas subject to industrial risk may have different economic trends than control areas.

Matching methods can then provide a solution as they rely on pairing each at-risk unit with *ax ante* similar control units. This approach can be combined with difference-in-difference and the ATT is then estimated by comparing the outcomes of each pair. Initially, matching estimation relied on finding pairs of observation having exactly the same *ex ante* observable characteristics. The main caveat of this approach is that one wants to include as many characteristics as possible in the matching process, but doing so reduces the chances of finding a twin observation for each treated one. Rosenbaum and Rubin, (1983) show that this "curse of dimensionality" can be resolved as long as the number of potential control variables is smaller than the set of observations, as it is equivalent to condition the outcomes on observable characteristics or on propensity score based on these characteristics. I adopt this now standard strategy in my analysis.

Following Smith and Todd, (2005), the difference-in-difference matching estimator can be written as follows :

$$\widehat{ATT} = \frac{1}{n_1} \sum_{i \in I_1 \cap S_p} \left[\left(Y_{it'}^{T=1} - Y_{it}^{T=1} \right) - \sum_{j \in I_0} w\left(i,j\right) \left(Y_{it'}^{T=0} - Y_{it}^{T=0} \right) \right]$$

where S_p is the support of the propensity score p, n_1 the number of locations in $I_1 \cap S_p$. The weights $w\left(i,j\right)$ depend on both the distance between p_i and p_j (the propensity score estimates for locations i and j), and the chosen estimation method (results shown in this paper use a *kernel* method with optimal bandwidths).

2.4 Defining at-risk and control locations

The location of dangerous plants is considered as public information, thus a government website[6] provides the complete list along with some information about each plant. The European Council "Seveso Directive" defines two levels of risk, depending on the potential damage that could occur and the estimated probability of such an occurrence. In this paper, the analysis is focused on upper-tier Seveso plants, i.e. only the most dangerous ones.

In January 2014, I identified 613 upper-tier Seveso plants in France. Figure 3 shows that their distribution across the country is highly heterogeneous, although no area is really free of industrial risk. I consider that a given location is exposed to industrial risk if it is within two kilometers of a dangerous plant. This definition is restrictive enough to reasonably expect risk perception in the area to be high (for example, people are likely to see it from their house or to drive close to it during their daily errands), and wide enough to ensure that there are indeed housing units within this range.

Defining control areas is slightly more complex. The shift in risk perception should not affect behaviors in control areas, and they should be as similar as possible to the at-risk areas. Given that the AZF accident caused damage on housing up to 7 kilometers away from the plant, it is plausible that there should be some impact on risk perception in areas that are less than seven kilometers away from a given dangerous plant. I thus exclude areas that are less than 7 kilometers away from a dangerous plants. In addition, I restrict control areas not to be further than 10 kilometer away from a dangerous plant. It ensures that treated and control areas share unobserved characteristics such as local amenities available to the inhabitants. Because the area covered by a wider disk is greater than the area covered by a small disk, the control group remains larger that the at-risk group and this restriction does not lead to a loss of power.

In addition, I exclude the urban area of Toulouse, where the AZF accident took place. Indeed, in this area the accident not only changed risk perception, but also affected the overall economic context. Part of the housing stock was damaged, and the local labor market could have been affected by the destruction of the plant.

A different approach to the definition of the control group could have been to select areas located in the vicinity of manufacturing plants that are not subject to the Seveso regulation. At first sight, the main advantage of this method is that the characteristics of local housing markets are likely to be similar in the at-risk locations and other manufacturing neighborhoods. However, this approach is highly impractical. Indeed, one would need to set a criterion defining similar but not risky industries. The type of industry would not be sufficient, as the Seveso regulation applies to a vast array of industries. The number of employees in the plant would not necessarily be a good

[6] http://www.installationsclassees.developpement-durable.gouv.fr

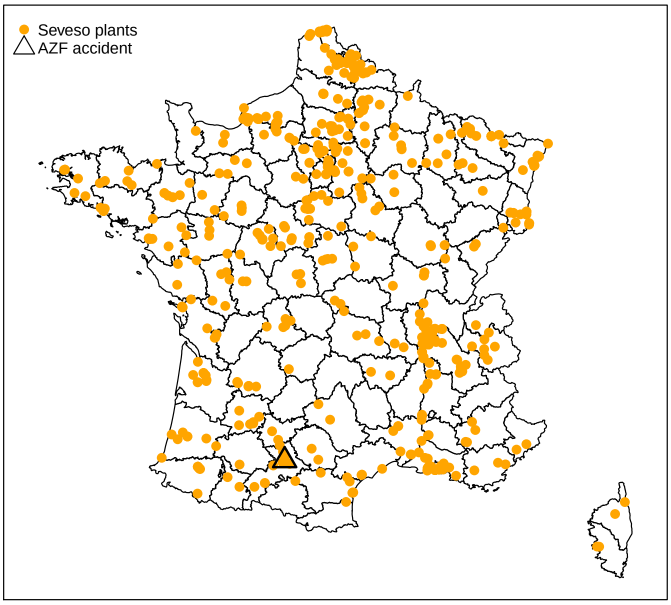


Figure 3: Distribution of upper-tier Seveso plants in France (January 2014)

criterion either, because there is a high heterogeneity in the size of Seveso plants. Selecting only plants with a high number of employees could in addition lead to selecting headquarters instead of actual production facilities. A third possibility would be to use a geographic information system to identify potential manufacturing areas, and select only those for which there is no identified Seveso plant. However, the Seveso plant census is not available in panel form, and there would thus be a risk of misclassifying former at-risk areas as control areas. Given these limitations, I believe that the approach chosen in this paper is more likely to yield unbiased results.

Even so, there is some discretion involved in the definition of at-risk and control areas and one concern could be that the results are driven by specific definitions. Tables A1 and A2 (in appendix) address this concern and show that the main results are however robust to slight changes in the definition of both groups. In addition, table A2 shows that reducing the size of the at-risk areas leads to stronger and more significant estimates. The estimated impact thus appears to be local and concentrated in the direct vicinity of Seveso plants.

3 Data sources and descriptive statistics of the panel

3.1 Housing and sales data

Data quality is one of the main assets of this paper. Among them are an exhaustive administrative database on housing in France (Filocom), and data collected at the local level on real estate transactions (Perval).

The Filocom database was created by the French tax administration using four different tax files on both housing units and households. As a result, this database includes households characteristics (for example age, earnings, family structure), housing characteristics (for example date of construction, square footage, several quality measures) and landlords characteristics for all 30 million housing units in France. The finest geographic scale that can be used to locate housing units are cadastral plan sections, which amount roughly to a block (I will refer to them as blocks in the remainder of this paper). These sections contain on average over 200 housing units but this measure can vary greatly depending on how urban or rural the section is. In particular, it is much higher in the Paris region where housing is much denser than in other parts of France.

This database can provide precise and detailed insight as to the structure and characteristics of a neighborhood, but it cannot account for housing prices. To study the impact of the accident on real estate prices, I thus use the Perval notaries database as an alternative source. Short of tax files, it is the most comprehensive source on real estate transactions in France. This data is collected form notaries' offices and contains mostly information on the estate being sold, along with partial characteristics of sellers and buyers. Blocks are again the most precise geographic unit that can be used to assess the location of the estates, and they can be used to match both data sets.

Only even years of both datasets are available, thus in this paper I mostly use two years of data: 2000 and 2002, a year before and a year after the accident. The year 1998 of Filocom is also used to better assess the resemblance between at-risk and control areas before the accident. However transaction files before 2000 could not be recovered, and the 1998 Filocom version does not include many relevant variables that are present in the the following years. To analyze the medium term impact of the shock, I also use the year 2004 in the last subsection of this paper.

For computational reasons, I aggregate both data sets at the block level. In particular, I compute the number of housing units per block, and I define the share of housing units that are vacant, used as principal or secondary housing, and leased as furnished housing. All these categories are built on tax definitions, and mostly rely on council tax exemptions. I also compute the share of dwellings built within two years (construction rate) or destroyed within two years (destruction rate), and the selling rate, defined as the number of transactions occurring in a given block over the number of housing units within this block. All these variables can be defined for any block.

In addition, I compute several indicators that can only be defined for a subset of the initial sample. First, by definition the average transaction prices (in euro per square meter) only exist for blocks in which there is at least one transaction. Second, for tax collection reasons the household composition is only known for dwellings used as principal housing. These variables can thus be computed at the block level but represent a smaller sample of dwellings. To study the evolution of household composition and wealth, I use the number of consumption units per household, the number of children under 6 and 18 years old per household and the annual earnings per consumption unit. These outcomes are averaged over blocks. I also compute the share of homeowners, private-market renters and social housing renters within each block. These variables are defined as shares of principal housing.

3.2 Descriptive statistics

This paper relies on the analysis of over 21,000 blocks, which account for more than a third of the total housing stock in France. 5,182 (a little over 20%) of them are subject to industrial risk. Table 1 shows that on average, they are significantly different from the control areas before the AZF accident. At-risk locations are denser, but tend to have a higher vacancy rate and a lower share of secondary housing. Both these elements could result from a lower overall level of amenities or a higher mobility in at-risk areas. However, even if the difference in two-year mobility rates between both groups is statistically significant, it seems too small to account fully for these composition differences. At-risk and control areas also have different household characteristics. Residents are less likely to be homeowners, and their average earnings are lower. Thus the populations living in both groups were also different before the accident. These *ex ante* differences suggest that even

Table 1: Descriptive statistics

Variables	Control (1)	Treated (2)	Diff. (1) - (2) (3)
Nb. Households	217	255	-38***
Construction rate	3.40%	3.17%	0.23**
Destruction rate	0.83%	1.09%	-0.26***
2-year mobility rate	19.5%	21.41%	-1.90***
Vacancy rate	7.42%	8.20%	-0.78***
Principal housing	85.36%	88.50%	-3.14***
Secondary housing	6.93%	3.06%	3.87***
Furnished housing	0.28%	0.24%	0.04
Transaction rate	1.45%	1.45%	0.00
Household size (consumption units)	1.73	1.72	0.02***
Nb. of children under 18 y.o. per household	0.59	0.58	0.01**
Nb. of children under 6 y.o. per household	0.19	0.19	0.00
Annual household earnings per consumption unit (€)	13483	11654	1829*
Owner-occupied	71.40%	62.45%	8.95***
Rented	17.54%	20.45%	-2.91***
Rented social housing	5.88%	12.41%	-6.53***
Price per square meter (€/m^2)	1111	977	134.86**
Houses price per square meter (€/m^2)	1124	1012	111.78**
Apartments price per square meter (€/m^2)	1275	986	289.82**
Nb. Obs. (blocks)	21,087	5,182	–

Source: Filocom, Perval. Author's calculations.

Note: This table presents average block values in the main sample.

Reading: ***: significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

before the AZF accident, overall housing markets and neighborhoods characteristics in both groups were not similar.

Table 2: Price subsample: selection effect

Variables	All (1)	Price panel (2)	Diff. (1) - (2) (3)
Nb. Households	224	368	-143
Construction rate	3.35%	2.97%	0.39***
Destruction rate	0.88%	0.93%	-0.05
Vacancy rate	7.57%	7.04%	0.53***
Principal housing	85.98%	87.87%	-1.88
Secondary housing	6.17%	4.79%	1.38***
Furnished housing	0.27%	0.30%	-0.03
Transaction rate	1.45%	2.65%	-1.20
Household size (consumption units)	1.73	1.69	0.04***
Nb. of children under 18 y.o. per household	0.59	0.58	0.01***
Nb. of children under 6 y.o. per household	0.19	0.19	0.00
Annual household earnings per consumption unit (€)	13122	13054	68
Owner-occupied	69.64%	64.88%	4.76***
Rented	18.12%	20.60%	-2.48
Rented social housing	7.17%	11.07%	-3.90
Price per square meter (€/m^2)	1081	1121	-39.97
Houses price per square meter (€/m^2)	1099	1146	-47.29
Apartments price per square meter (€/m^2)	1196	1205	-9.02
Nb. Obs. (blocks)	26,269	10,223	–

Source: Filocom, Perval. Author's calculations.

Note: The price panel is a subset of observations for which transactions occurred both before and after the accident, a necessary condition to study its impact on transaction prices. This table presents average block values in the price panel.

Reading: *** : significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

These descriptive statistics thus confirm that a direct comparison of at-risk and control areas cannot provide a causal estimation of the impact of risk perception. Moreover, figure A1 (in appendix) shows that the common trend assumption is not reliable for at least some of the outcomes before the accident.[7]

Additionally, the difference-in-difference methodology implies an important restriction for the analysis of prices. Indeed, the average transaction price can only be computed for blocks with a non-zero number of transactions. The effect of the accident on prices can thus only be computed for blocks where transactions happened both in 2000 and in 2002. As table 2 shows, this subsample is about twice as small as the main sample, and its characteristics strongly and significantly differ. The conclusion drawn on this subsample can then suffer from selection bias. As such, they cannot necessarily be extended to the full sample.

[7] More formal tests of this necessary condition can be found for available variables in the next section (table 4).

4 Findings and discussion

4.1 Propensity score estimation

Propensity score matching is generally estimated in a two-step procedure. The first one is the estimation of the propensity score model: $\mathbb{P}[T = 1|X]$.

This estimation is conventionally done with a Probit model. I include in the estimation both variables describing the blocks in 2000 and trends of these variables between 1998 and 2000. More precisely, the propensity score is estimated using *ex ante* area characteristics (rural or urban area dummies, number of households, number of sales in 2000, share of collective housing, rented housing units, social housing, vacant housing units, main or secondary residence), *ex ante* household characteristics (number of persons, number of children under 18 and 6 y.o., average income and income per consumption unit) and *ex ante* trends for these variables (when available in the data). A separate estimation is done for each subsample, but the overall quality of the estimation does not vary much accross the different populations. Indeed, I perform the test introduced by Dehejia and Wahba, (2002) and present the results in table 3. This test checks whether conditioning the observations by the score indeed balances the covariates between the at-risk and control groups. To do so, one has to first stratify observations in order to verify $\hat{p}(X) \perp T$. In a second step, one checks that the observable characteristics are not significantly different for at-risk and control observations in each strata. The score can then be considered balanced if the rejection rate at the 5% level does not exceed 5%. Table 3 shows that this is the case for the main six samples.

Table 3: Dehejia and Wahba balancing test

	All	South-West quarter	Chemical plant
Main panel	4.30%	3.10%	3.78%
Price subsample	2.14%	1.15%	4.05%
Number of blocks	26,269	3,530	8,191

Source: Filocom, Perval. Author's calculations.

Reading: 4.30% of balancing conditions are rejected in the main panel.

In addition, table 4 shows the estimated differences-in-differences estimates for the period 2000-1998. Only a few outcomes are reported in this table, as all variables are not available in the 1998 version of the data. The results however are reassuring regarding the quality of the matching procedure. Columns 1 and 2 show the results obtained with a simple differences-in-differences, while columns 3 and 4 show the results obtained by combining propensity score matching with this standard approach. As expected, several differences are rather large in magnitude and significantly different from zero in the first part of this table. This illustrates the initial differences in trends between at-risk and control areas. However, the second part of the table shows that the matching procedure greatly reduces the magnitude of the estimated difference, and that they are no longer significant at the standard levels of 1%, 5% or 10%.

4.2 Main results

The main results can be found in the first column of tables 5, 6 and 7. Columns 2 and 3 of these tables show the results for two sub-samples, that are detailed in the next subsection. Table 5 shows that on average, the transaction prices in the at-risk areas decrease by 23 euros per square meter (about 2% of the initial price) after the accident (compared with the control areas). The effect is stronger for apartments than for houses, and it is not significant for houses alone. Even

Table 4: Placebo tests: "effect" between 1998 and 2000

	Unmatched DiD estimate		Matched DiD estimate	
	ATT	St. error	ATT	St. error
Vacancy rate (pp)	0.140*	0.078	0.055	0.083
Owner-occupied (pp)	0.084	0.088	0.025	0.087
Rented (pp)	-0.351***	0.095	-0.031	0.105
Rented social housing (pp)	0.171***	0.060	0.009	0.066
Household size (number of people)	0.000	0.003	-0.001	0.003
Number of children <18 y.o.	0.000	0.002	-0.001	0.002
Total annual household earnings (€)	-742***	191	-73	150
Annual household earnings per consumption unit (€)	-465***	117	-47	88
Number of blocks	26,269		26,269	

Source: Filocom, Perval. Author's calculations.

Sample : treated and control blocks.

Note: the first column shows results estimated with a simple difference-in-difference, while the second column shows results obtained through propensity score kernel matching with optimal bandwidth.

Reading : *** : significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

in the short term, the accident thus seems to have a rather strong impact on transaction prices, in particular for apartments. The difference between houses and apartments may come from the fact that houses are even more multi-dimensional than apartments, as the price of a house notable includes the price of its land. The intensity of the price decrease is at odds with the findings of Grislain-Letrémy and Katossky, (2013), which is the only other attempt at studying the impact of industrial risk on housing markets in France. They found that the housing prices in the vicinity of three hazardous industrial plants (Rouen, Dunkirk, Bordeaux) reflect the negative and positive amenities of the plants[8] but that those price differences were not modified by the AZF accident. However, it is possible that there should not be a significant impact in the vicinity of these three particular sites but that the accident should on average have changed the vicinity of hazardous plants.

Moreover, table 6 shows that the price effect hides other significant changes on the housing markets. Indeed, the vacancy rate increases in the at-risk areas, in relative terms, compared to the control areas. In fact, figure A1 shows that this relative increase is in reality a smaller decrease than in the control areas. This does not intrinsically invalidate the interpretation that all else held equal, the vacancy rate in at-risk areas would have been lower in the absence of an information shock. However, it is possible that part of the decrease in vacancies in the control group was due to people living the at-risk areas for control areas. Such displacement effects would bias the estimates presented in this paper and increase their magnitude. Although I cannot completely rule this explanation out, there is evidence in Bléhaut and Mauroux (2016)that such effects did not happen in the wake of other information shocks between 2003 and 2012.

As expected, the share of principal housing decreases conversely to this relative increase of vacancies. All in all, I find a significant increase in the share of vacant houses by 0.3 percentage points (pp), that seems to be fully compensated by a decrease in the share of principal housing. In magnitude, the effect on the vacancy rate is roughly equivalent to the one on price and amounts to a 3% increase. The selling rate also decreases a little, although it is not significant at the 10% level. These changes are consistent with a process in which housing units that are put on the

[8]For instance the Seveso plant near Bordeaux is a former gunpowder factory, surrounded by a park and the neighborhoods around then plant are green and quiet, so contrary to other industrial areas there are high positive amenities in this location.

Table 5: Matching estimation – sales characteristics

	All	South-West quarter	Chemical plants
Price per square meter (€/m^2)	-23.347**	-60.614*	-29.095*
Houses price per square meter (€/m^2)	-11.123	-48.517	-20.933
Apartments price per square meter (€/m^2)	-34.074***	-163.341***	-30.166
Number of blocks	10,223	1,097	3,185
Number of housing units (2000)	3,760,775	387,359	1,104,001
Number of sales (2000)	86,836	7,419	23,928

Source : Filocom, Perval. Author's calculations.

Sample : treated and control blocks (all types of housing units).

Note : the results are obtained through propensity score kernel matching with optimal bandwidth. The propensity score is estimated separately for each column.

Reading : *** : significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

Table 6: Matching estimation – housing stock evolution

	All	South-Ouest quarter	Chemical plant
Construction rate (pp)	0.130	0.399	-0.155
Destruction rate (pp)	0.036	0.033	-0.034
Vacancy rate (pp)	0.290***	0.528**	0.632***
Principal housing rate (pp)	-0.340***	-0.743***	-0.600***
Secondary housing rate (pp)	0.050	0.239	-0.031
Furnished housing rate (pp)	-0.001	-0.024	-0.001
Selling rate (pp)	-0.067	-0.258**	-0.085
Number of blocks	26,269	3,530	8,191
Number of housing units (2000)	5,896,135	526,550	1,469,465

Source : Filocom, Perval. Author's calculations.

Sample : treated and control blocks (all types of housing units).

Note : the results are obtained through propensity score kernel matching with optimal bandwidth. The propensity score is estimated separately for each column.

Reading : *** : significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

Table 7: Matching estimation – household characteristics

	All	South-Ouest quarter	Chemical plant
Household size (consumption units)	-0.003**	-0.003	0.002
Nb of children < 18 y.o.	0.000	-0.006	0.005
Nb of children < 6 y.o.	-0.002	0.001	0.001
Household earnings per consumption unit (€)	-235*	-265	-213
Owner-occupied (pp)	0.122	0.678**	0.082
Rented (pp)	-0.067	-0.490	-0.125
Rented social housing (pp)	0.016	-0.169*	0.064
Number of blocks	26,269	3,530	8,191
Number of housing units (2000)	5,080,150	430,105	1,275,705

Source : Filocom, Perval. Author's calculations.

Sample : treated and control blocks (principal housing only).

Note : the results are obtained through propensity score kernel matching with optimal bandwidth. The propensity score is estimated separately for each column.

Reading : *** : significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

market are not sold or rented as quickly in the at-risk areas as in the control areas. In particular, it is consistent with the disposition effect illustrated on the Boston housing market (Genesove and Mayer, 2001): sellers are reluctant to lower their expected price, while buyers preferences have already reacted to the information shock caused by the accident. As a result, in the short term less transactions happen. As expected in this short term perspective, there is however no effect of the accident on the construction and destruction rates.

Finally, table 7 shows the estimated impact of the accident on the main household characteristics (these estimations only include principal housing). Most of them remain stable, and in particular there is no significant change in the average number of children per household. This result is at odds with the hypothesis that families would react more to environmental quality and risk, but it could also result from the higher mobility cost these households face. Both effects could cancel out. The average household earnings tend to decrease in the at-risk areas. This effect is not due to household composition, since I consider households earnings per consumption units. They decrease by 250 euro per consumption unit over the 2000-2002 period. This relative pauperization can indicate that some residential sorting based on industrial risk is happening. It is all the more likely as poorer households face higher moving costs and are less likely to outbid others. One should however note that this effect is not highly significant.

As noted previously, these main results are remarkably robust to slight changes in the definition of control or at-risk locations. Indeed, table A1 shows that these conclusions mostly hold for six alternative control groups, with the notable exception of the earnings effect. This effect indeed varies from -112 euros per consumption unit to -235 (the highest is measured with the initial control group). In addition, there is a lack of statistical significance with two of the control groups. Furthermore, table A1 highlights the local nature of the consequences of the information shock. All the effects are higher in magnitude and tend to be more significant with more restrictive definitions of the at-risk areas.

One could wonder whether the matching procedure played an important role in these results. Tables A3 to A5 in appendix allow to compare the simple differences-in-differences results with the matching estimates. In almost all cases, the former are greater in magnitude and tend to be more significant. For example, the simple differences-in-differences estimate for the impact of the accident on average apartments price is almost twice as big as that of matching, and significant at the 1% level. This pattern also appears with various degrees of intensity in table A4 and A5, which is consistent with the existence of a bias due to the violation of the common trend assumption. There is one notable exception to this conclusion. Indeed, the annual household earnings in the unmatched specification is both smaller and not significant, which casts a doubt on the validity of this result.

4.3 Is the effect stronger when the shift in risk perception is higher?

Tables 5, 6 and 7 also show the estimated effect on two sub-samples in which the shift in risk perception is likely to be higher. Both these sub-samples are represented on figure 4. The first one is the South-West quarter of France, where the accident originally happened. Indeed, local press and television channels are likely to have reported more heavily and for a longer period of time on the subject. Local population is thus likely to have been more exposed to the information than the rest of France. Moreover, people living in this part of the country could have felt more concerned with an accident that hit a big city closer to home. The second sub-sample is based on the type of industry: the AZF plant was a chemical factory. I thus isolate the neighborhoods of chemical Seveso plants, although this is a rather heterogeneous group both in terms of location and in terms of type of plants. Indeed, chemistry is a very large denomination that is likely to cover plants with very diverse aspects.

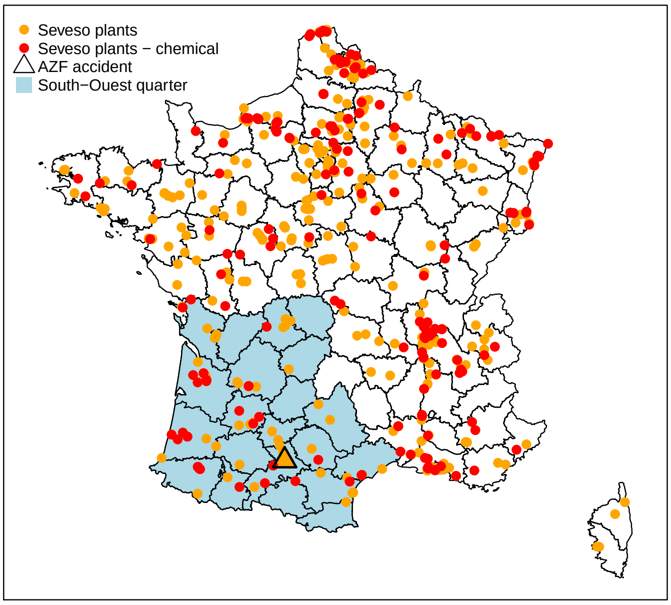


Figure 4: Risk areas subsamples

Overall, the estimated impacts indeed tend to be stronger in these two sub-sample than in the general population, especially in the South-West quarter. Table 5 shows the variation for price estimates in both sub-samples. Similarly, the point estimates of price in the South-West quarter and around chemical plants is stronger. There is a lack of significance in the chemical plants sub-sample that is consistent with the heterogeneity of this group, while the overall impact in the South-West quarter is estimated to be about 160 euros per square meter for apartments and is highly significant. One should use this result with caution however, since the sample is quite small and excludes Toulouse, one of the biggest metropolitan area of this part of France.

In addition, table 6 shows that the effect on the vacancy rate and share of principal housing are at least twice as high in both sub-samples. Moreover, the selling rate decreases more sharply in the South-West quarter than in the main panel, which is consistent with a stronger negative demand shock. Table 7 also show that some effects vary differently in both sub-samples. Indeed, the effect on household earnings is higher in the South-West quarter, although it is too imprecise to be statistically significant. More importantly, the type of occupancy is significantly affected in the South-West quarter, contrary to other samples. Indeed, there is a statistically significant 0.6 percentage point increase in the share of Homeowners in this area, compensated by an decrease in the share of renters both in the private and social markets (although only the latter is significant). This change is consistent with homeowners facing higher relocation costs than renters, since they often have to sell their property before moving. In a context of disposition effect, they are likely to wait even longer before affording to move.

4.4 How does the effect evolve over time?

Tables 8 and 9 compare the short-time estimates (for the 2000-2002 period) with the estimates of the following period (between 2002 and 2004) for price outcomes and housing stock outcomes. On the one hand, table 8 shows that the negative impact of the information shock on transaction prices increases in magnitude in the second period. Far from diminishing over time, the estimates are twice as high for this period, and significant at the 1% level both for houses and apartments. The decrease in prices between 2002 and 2004 in the at-risk areas is 55 euros per square meter compared with control areas. Table 9 on the other hand shows that there is no additional impact on the vacancy rate. On the contrary, there seems to be an intensification of the volume of transactions in the at-risk areas with respect to control areas. Indeed, we find a significant 0.15 pp increase in the selling rate between 2002 and 2004.

Table 8: Matching estimation – medium-term effect – sales characteristics

	2000-2002	2002-2004
Price per square meter (€/m^2)	-23.347**	-53.457***
Houses price per square meter (€/m^2)	-11.123	-61.109***
Apartments price per square meter (€/m^2)	-34.074***	-60.024***
Number of blocks	10,223	10,332
Number of housing units (2000 and 2002)	3,760,775	3,706,636
Number of sales (2000 and 2002)	86,836	84,411

Source : Filocom, Perval. Author's calculations.

Sample : at-risk and control blocks (all types of housing units).

Note : the results are obtained through propensity score kernel matching with optimal bandwidth.

Reading : *** : significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

Additional confounding factors could have arised between the AZF accident and 2004, thus the interpretation of these findings is subject to more caution than our main results. In particular, a new law regarding industrial risk prevention was passed in 2003. However, this exercise can lead to insights regarding the rationality of some of the behaviors observed over the 2000-2002 period. In particular, it would be rational not to sell one's property following the AZF information shock if one correctly anticipates market prices to go back to their previous value. This does not however seem to be the case. In light of this evolution, the intensification of the volume of transactions between 2002 and 2004 could be driven both by a price effect (more potential buyers willing to purchase a property at a lower cost) or a shift in supply (more sellers willing to sell their property

Table 9: Matching estimation – medium-term effect – housing stock evolution

	2000-2002	2002-2004
Construction rate (pp)	0.130	-0.116
Destruction rate (pp)	0.036	-0.052
Vacancy rate (pp)	0.290***	-0.027
Principal housing rate (pp)	-0.340***	0.103
Secondary housing rate (pp)	0.050	-0.067
Furnished housing rate (pp)	-0.001	-0.010
Selling rate (pp)	-0.067	0.150***
Number of blocks	26,269	26,269
Number of housing units (2000 and 2002)	5,896,135	5,977,117

Source : Filocom, Perval. Author's calculations.

Sample : at-risk and control blocks (all types of housing units).

Note : the results are obtained through propensity score kernel matching with optimal bandwidth.

Reading : *** : significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

at a lower cost once it is established that the downward shift in price is a lasting one). The second period estimates are thus still consistent with the disposition effect, but question the validity of sellers' anticipations.

4.5 Is the effect driven by homeowners?

Tables A6 to A8 (in appendix) show the estimated impact of the AZF accident in the sample of blocks with an initially low share of homeowners. The cut-off used to define high and low shares of homeowners is the median value for at-risk blocks in 2000.

The most striking feature of these tables is that the whole effect displayed in the main specification is driven by this sub-sample. All the results for blocks with an initially low share of homeowners are thus higher in magnitude, and generally more significant than in the main specification. In particular, as expected, the negative impact on the vacancy rate in at-risk areas and positive impact on principal housing are reinforced in this sub-sample (table A7). There is however no impact of the accident on the share of homeowners (table A8). Finally, the reduction in price is entirely driven by these blocks.[9]

How can these results be interpreted? First, the higher impact on vacancy and neighborhood composition can be explained by the relatively lower mobility costs of renters with respect to homeowners. In areas with a low share of homeowners, renters are likely to be more numerous. Second, in these neighborhoods observed transactions are more likely to be buy-to-rent than buy-to-live, compared with areas with a high share of homeowners. One can thus deduce that the price effect is mostly driven by buyers anticipating lower expected rent and higher search cost to find a tenant at least in the short term. Both of these expectations amount to a lower return to investment, and would thus decrease prices.

4.6 Is the effect homogeneous with initial price?

Tables A9 to A11 (in appendix) show the estimated impact of the AZF accident depending on the initial level of prices. The cut-off used to define high and low initial prices is the median value for at-risk blocks in 2000.

[9]Additional results, not reported here for the sake of clarity, show that there is no significant impact on any of the selected outcomes for the sample of blocks with an initially high share of homeowners.

For blocks with an initially low level of transaction prices, the price decrease seems to be stronger than in the overall sample. In particular, it is roughly twice as strong as in the overall sample and for apartments and houses taken separately, and it is significant for both. For blocks with an initially high transaction price, the effect is much lower and becomes insignificant. However, the contrary happens for vacancy: there thus seems to be a trade-off between vacancy and price reduction. There are two potential explanation for this. First, it is possible that there would be higher liquidity constraints for homeowners wishing to leave low-price areas since they are likely to have a lower capital. They are thus constrained to sell more rapidly to afford to move, even if it means lowering their expected price. second, there may be some unobserved heterogeneity in amenities. Indeed, Grislain-Letrémy and Katossky, (2013) showed that in the well-off neighborhood of a gun-powder Seveso plant in Bordeaux, positive local amenities completely dominate the risk effect.

Finally, the change in average household earnings is mostly driven by areas with an initially low price. This does not, however, invalidate the hypothesis that richer households are more mobile than poorer ones. Indeed, within these blocks it may very well be that richer people are more likely to leave. Conversely, it could also be that poorer household tend to move in. It is however impossible to answer this question with the aggregation strategy used in this paper is left open for further research.

5 Conclusion

This paper contributes to the literature on housing markets and environmental quality. The main contribution is to show empirically that housing markets react to symmetric information shocks that should have no impact in a perfect information setting. In practice, hedonic analysis of environmental goods relies on the assumption that both buyers and sellers have a full knowledge of the environmental quality of their house. This paper provides evidence that it may not always be valid, as in many settings residents are likely to have a biased risk perception. Indeed, had an hedonic analysis of willingness to pay to avoid industrial risk been undertaken before the AZF disaster, one would have concluded to a very small impact of this characteristic. After the accident however, the differentiation between at-risk and similar safer neighborhoods durably reinforces. Taking the imperfect information into account seems to be crucial in understanding the traditionally low price-elasticity of housing with respect to environmental quality estimated in the literature. Moreover on the short term, the relation between price and environmental quality can hide more complex mechanisms on the housing markets. From a policy perspective, this paper gives credit to measures such as mandatory assessment and disclosure in the case of housing transactions.

The quasi-experimental setting provided by the AZF accident strengthens the credibility of the results, as they are less likely to be driven by other changes. Given the scope of the shock on risk perception studied in this case, it is not surprising that I should find evidence supporting a rather strong impact of risk perception on housing prices. However, I also find that at-risk neighborhoods react to industrial risk perception in other complementary ways. In particular in the short term, the vacancy rate increases in at-risk neighborhoods and the transaction rate decreases slightly. In addition, households living in a dangerous area after the shift in risk perception tend to be poorer, which is consistent with residential sorting. There is however no evidence of families being more sensitive to the additional information on risk, which may result from higher mobility costs. Moreover, the effects tend to be stronger in more exposed neighborhoods, and are mostly driven by areas with an initially low share of homeowners. In the medium term, the price effect tends to grow stronger and is associated with an increase in the volume of transaction, which questions the rationality of the strategy adopted by homeowners in the short term and shows durable differentiation of the housing market based on industrial risk.

References

Case, K. E. and R. J. Shiller (1989). “The Efficiency of the Market for Single-Family Homes”. *The American Economic Review* 79.1, pp. 125–137.

Chabbal, J. (2005). “Le risque invisible”. *Politix* 2, pp. 169–195.

Chay, K. and M. Greenstone (2005). “Does Air Quality Matter? Evidence from the Housing Market”. *Journal of Political Economy* 113.2, pp. 376–424.

Currie, J. et al. (2015). “Environmental health risks and housing values: evidence from 1,600 toxic plant openings and closings”. *The American economic review* 105.2, pp. 678–709.

Davis, L. C. (2011). “The effect of Power Plants on Local Housing Values and Rents”. *The Review of Economics and Statistics* 93.4, pp. 1391–1402.

Davis, L. W. (2004). “The effect of health risk on housing values: Evidence from a cancer cluster”. *The American Economic Review* 94.5, pp. 1693–1704.

Dehejia, R. H. and S. Wahba (2002). “Propensity score-matching methods for nonexperimental causal studies”. *Review of Economics and Statistics*, pp. 151–161.

Duflo, E. and E. Saez (2003). “The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment”. *The Quarterly Journal of Economics* 118.3, pp. 815–842.

Gayer, T., J. T. Hamilton, and W. K. Viscusi (2000). “Private Values of Risk Tradeoffs at Superfund Sites: Housing Market Evidence on Learning about Risk”. *The Review of Economics and Statistics* 83.3, pp. 439–451.

Genesove, D. and C. Mayer (2001). “Loss Aversion and Seller Behavior: Evidence from the Housing market”. *The Quarterly Journal of Economics* 116.4, pp. 1233–1260.

Glaeser, E. L. and J. Gyourko (2007). *Arbitrage in housing markets.* Tech. rep. National Bureau of Economic Research.

Grainger, C. A. (2012). “The distributional effects of pollution regulations: Do renters fully pay for cleaner air?” *Journal of Public Economics* 96.9, pp. 840–852.

Greenstone, M. (2002). “The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census Manufactures”. *Journal of Political Economy* 110.6, pp. 1175–1219.

Greenstone, M. and J. Gallagher (2008). “Does Hazardous Waste Matter ? Evidence from the Housing Market and the Superfund Program”. *Quarterly Journal of Economics*.

Grislain-Letrémy, C. and A. Katossky (2013). “Les risques industriels et le prix des logements”. *Economie et Statistique* 460-461, pp. 79–106.

Hastings, J. S. and J. M. Weinstein (2008). “Information, School Choice, and Academic Achievement: Evidence from Two Experiments”. *The Quarterly Journal of Economics*, pp. 1373–1414.

Jensen, R. (2010). “The (perceived) returns to education and the demand for schooling”. *Quarterly Journal of Economics* 125.2.

Kiel, K. A. and M. Williams (2007). “The Impact of Superfund Sites on Local Property Values: Are all Sites the Same?” *Journal of Urban Economics* 61, pp. 170–191.

Muehlenbachs, L., E. Spiller, and C. Timmins (2015). “The housing market impacts of shale gas development”. *The American Economic Review* 105.12, pp. 3633–3659.

Parmeter, C. F. and J. C. Pope (2012). “Quasi-Experiments and Hedonic Property Value Methods”.

Pope, J. C. (2008). “Do seller disclosures affect property values? Buyer information and the hedonic model”. *Land Economics* 84.4, pp. 551–572.

Rosen, S. (1974). “Hedonic prices and implicit markets: product differentiation in pure competition”. *Journal of political economy* 82.1, pp. 34–55.

Rosenbaum, P. R. and D. B. Rubin (1983). “The Central Role of the Propensity Score in Observational Studies for Causal Effects”. *Biometrika* 70, pp. 41–55.

Rubin, D. (1974). "Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies". *Journal of Educational Psychology* 66, pp. 688–701.
Shefrin, H. and M. Statman (1985). "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence". *The Journal of Finance* XL.3, pp. 777–790.
Smith, J. A. and P. E. Todd (2005). "Does matching overcome LaLonde's critique of nonexperimental estimators?" *Journal of Econometrics* 125.1-2, pp. 305–353.
Viscusi, W. K. and J. T. Hamilton (1999). "Are Risk Regulators Rational? Evidence from Hazardous Waste Cleanup Decisions". *American Economic Review* 89.4.
Viscusi, W. K. (1990). "Do smokers underestimate risks?" *Journal of political Economy*, pp. 1253–1269.

Appendix

Table A1: Matching estimation – robustness to the size of the control group

	Ref. (7-10 km)	7-11 km	7-12 km	7-13 km	8-11 km	9-12 km	10-13 km
Price per square meter (€/m^2)	-23.347**	-20.513**	-22.327**	-20.836**	-17.255*	-19.752*	-13.666
Houses price per square meter (€/m^2)	-11.123	-10.996	-12.447	-11.562	-8.247	-11.459	-10.795
Apartments price per square meter (€/m^2)	-34.074***	-32.054***	-35.454***	-36.108***	-28.832**	-31.911**	-34.438**
Construction rate (pp)	0.130	0.134	0.056	0.092	0.219	0.219	0.200
Destruction rate (pp)	0.036	0.011	-0.027	-0.043	0.009	-0.096	-0.189***
Vacancy rate (pp)	0.290***	0.279***	0.286***	0.277***	0.318***	0.320***	0.313***
Principal housing rate (pp)	-0.340***	-0.316***	-0.329***	-0.325***	-0.334***	-0.321***	-0.294***
Secondary housing rate (pp)	0.050	0.041	0.048	0.052	0.022	0.010	-0.013
Furnished housing rate (pp)	-0.001	-0.003	-0.005	-0.004	-0.005	-0.008	-0.006
Selling rate (pp)	-0.067	-0.078*	-0.075*	-0.074*	-0.092*	-0.101**	-0.105**
Household size (consumption units)	-0.003**	-0.004**	-0.003**	-0.003**	-0.004**	-0.004**	-0.003*
Nb of children $<$ 18 y.o. per household	0.000	-0.001	-0.001	-0.001	-0.001	-0.002	-0.001
Nb of children $<$ 6 y.o. per household	-0.002	-0.003*	-0.002*	-0.003**	-0.003**	-0.003*	-0.003**
Annual household earnings per consumption unit (€)	-235*	-205**	-176*	-164**	-223*	-162	-112
Owner-occupied (pp)	0.122	0.109	0.132	0.127	0.140	0.127	0.130
Rented (pp)	-0.067	-0.027	-0.091	-0.085	0.015	-0.054	-0.050
Rented social housing (pp)	0.016	-0.015	0.016	0.018	-0.065	0.003	-0.005

Source : Filocom, Perval. Author's calculations.

Sample : blocks within a 13km radius around Seveso plants (all types of housing units).

Note : the results are obtained through propensity score kernel matching with optimal bandwidth.

Reading : *** : significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

Table A2: Matching estimation – robustness to the size of the at-risk group

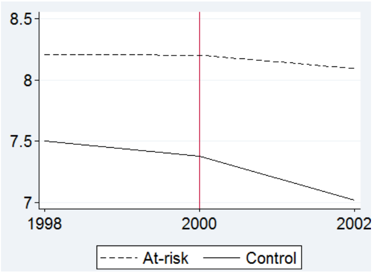
	0-1800 m	Ref. (0-2000 m)	0-2200 m	0-2400 m
Price per square meter (€/m^2)	-25.209***	-23.347**	-22.398**	-20.754**
Houses price per square meter (€/m^2)	-11.594	-11.123	-7.772	-7.869
Apartments price per square meter (€/m^2)	-40.194***	-34.074***	-31.992***	-26.118**
Construction rate (pp)	0.183	0.130	0.092	0.150
Destruction rate (pp)	0.045	0.036	0.032	0.009
Vacancy rate (pp)	0.307***	0.290***	0.260***	0.219***
Principal housing rate (pp)	-0.397***	-0.340***	-0.309***	-0.259***
Secondary housing rate (pp)	0.074	0.050	0.051	0.044
Furnished housing rate (pp)	0.016	-0.001	-0.003	-0.004
Selling rate (pp)	-0.058	-0.067	-0.039	-0.025
Household size (consumption units)	-0.003*	-0.003**	-0.003*	-0.003*
Nb. of children < 18 y.o. per household	0.001	0.000	0.000	0.000
Nb. of children < 6 y.o. per household	-0.002	-0.002	-0.002	-0.002
Annual household earnings per consumption unit (€)	-237*	-235*	-231*	-205*
Owner-occupied (pp)	0.138	0.122	0.104	0.087
Rented (pp)	-0.007	-0.067	-0.060	-0.051
Rented social housing (pp)	-0.015	0.016	0.010	0.015

Source : Filocom, Perval. Author's calculations.

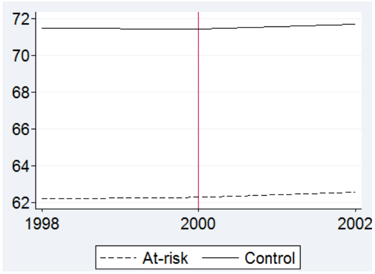
Sample : blocks within a 10km radius around Seveso plants (all types of housing units).

Note : the results are obtained through propensity score kernel matching with optimal bandwidth.

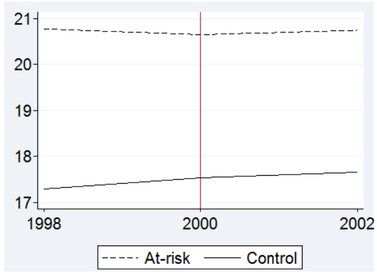
Reading : *** : significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.



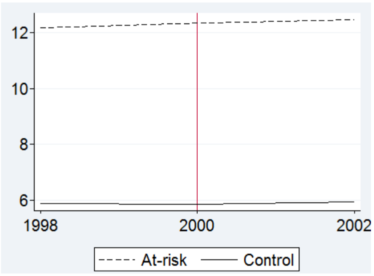
(a) % Vacancy



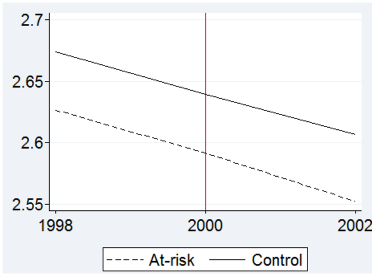
(b) % Owner-occupied



(c) % Rented

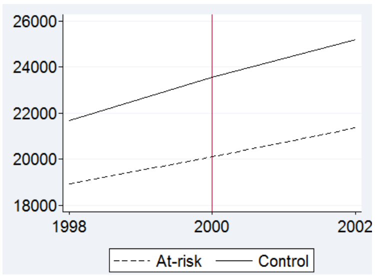


(d) % Social housing



(e) Number of people per household

(f) Number of children per household



(g) Total annual earnings per household (€)

Figure A1: Outcome trends before and after the AZF accident

Table A3: Specification robustness – sales characteristics

	Simple DiD	Matching DiD
Price per square meter (€/m^2)	-33.248***	-23.347**
Houses price per square meter (€/m^2)	-19.631	-11.123
Apartments price per square meter (€/m^2)	-56.398***	-34.074***
Number of blocks	10,223	10,223
Number of housing units (2000)	3,760,775	3,760,775
Number of sales (2000)	86,836	86,836

Source: Filocom, Perval. Author's calculations.

Sample: treated and control blocks (all types of housing units).

Note: the first column shows results estimated with a simple difference-in-difference, while the second column shows results obtained through propensity score kernel matching with optimal bandwidth.

Reading: DiD: differences-in-differences; ***: significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

Table A4: Specification robustness – housing stock evolution

	Simple DiD	Matching DiD
Construction rate (pp)	-0.035	0.130
Destruction rate (pp)	-0.049	0.036
Vacancy rate (pp)	0.234***	0.290***
Principal housing rate (pp)	-0.463***	-0.340***
Secondary housing rate (pp)	0.233***	0.050
Furnished housing rate (pp)	-0.005	-0.001
Selling rate (pp)	-0.026	-0.067
Number of blocks	26,269	26,269
Number of housing units (2000)	5,896,135	5,896,135

Source : Filocom, Perval. Author's calculations.

Sample : treated and control blocks (all types of housing units).

Note: the first column shows results estimated with a simple difference-in-difference, while the second column shows results obtained through propensity score kernel matching with optimal bandwidth.

Reading: DiD: differences-in-differences; ***: significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

Table A5: Specification robustness – household characteristics

	Simple DiD	Matching DiD
Household size (consumption units)	-0.003*	-0.003**
Nb of children < 18 y.o. per household	0.000	0.000
Nb of children < 6 y.o. per household	-0.002	-0.002
Annual household earnings per consumption unit (€)	-148	-235*
Owner-occupied (pp)	0.055	0.122
Rented (pp)	-0.076	-0.067
Rented social housing (pp)	0.032	0.016
Number of blocks	26,269	26,269
Number of housing units (2000)	5,080,150	5,080,150

Source : Filocom, Perval. Author's calculations.

Sample : treated and control blocks (principal housing only).

Note: the first column shows results estimated with a simple difference-in-difference, while the second column shows results obtained through propensity score kernel matching with optimal bandwidth.

Reading: DiD: differences-in-differences; ***: significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

Table A6: Heterogeneity with respect to the share of homeowners – sales characteristics

	All	Low homeowner share
Price per square meter (€/m^2)	-23.347**	-25.75**
Houses price per square meter (€/m^2)	-11.123	-6.619
Apartments price per square meter (€/m^2)	-34.074***	-33.836**
Number of blocks	10,223	4,792
Number of housing units (2000)	3,760,775	2,825,949
Number of sales (2000)	86,836	62,062

Source : Filocom, Perval. Author's calculations.

Sample : treated and control blocks (all types of housing units).

Note : the results are obtained through propensity score kernel matching with optimal bandwidth. The propensity score is estimated separately for each column.

Reading : *** : significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

Table A7: Heterogeneity with respect to the share of homeowners – housing stock evolution

	All	Low homeowner share
Construction rate (pp)	0.130	0.298
Destruction rate (pp)	0.036	-0.01
Vacancy rate (pp)	0.290***	0.418***
Principal housing rate (pp)	-0.340***	-0.490***
Secondary housing rate (pp)	0.050	0.076
Furnished housing rate (pp)	-0.001	-0.003
Selling rate (pp)	-0.067	-0.026
Number of blocks	26,269	9,622
Number of housing units (2000)	5,896,135	4,422,318

Source : Filocom, Perval. Author's calculations.

Sample : treated and control blocks (all types of housing units).

Note : the results are obtained through propensity score kernel matching with optimal bandwidth. The propensity score is estimated separately for each column.

Reading : *** : significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

Table A8: Heterogeneity with respect to the share of homeowners – household characteristics

	All	Low homeowner share
Household size (consumption units)	-0.003**	-0.007***
Nb of children < 18 y.o. per household	0.000	-0.001
Nb of children < 6 y.o. per household	-0.002	-0.003
Annual household earnings per consumption unit (€)	-235*	-328
Owner-occupied (pp)	0.122	-0.008
Rented (pp)	-0.067	0.007
Rented social housing (pp)	0.016	0.091
Number of blocks	26,269	9,622
Number of housing units (2000)	5,080,150	3,769,404

Source : Filocom, Perval. Author's calculations.

Sample : treated and control blocks (principal housing only).

Note : the results are obtained through propensity score kernel matching with optimal bandwidth. The propensity score is estimated separately for each column.

Reading : *** : significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

Table A9: Heterogeneity with respect to the initial level of price – sales characteristics

	All	Low initial price	High initial price
Price per square meter (€/m^2)	-23.347**	-26.818**	-26.446*
Houses price per square meter (€/m^2)	-11.123	-24.458*	-9.716
Apartments price per square meter (€/m^2)	-34.074***	-49.603***	-14.265
Number of blocks	10,223	3,955	6114
Number of housing units (2000)	3,760,775	1,307,498	2,435,533
Number of sales (2000)	86,836	24,166	62,474

Source : Filocom, Perval. Author's calculations.

Sample : treated and control blocks (all types of housing units).

Note : the results are obtained through propensity score kernel matching with optimal bandwidth. The propensity score is estimated separately for each column.

Reading : *** : significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

Table A10: Heterogeneity with respect to the initial level of price – housing stock evolution

	All	Low initial price	High initial price
Construction rate (pp)	0.130	-0.125	0.524
Destruction rate (pp)	0.036	-0.012	-0.027
Vacancy rate (pp)	0.290***	0.142	0.148
Principal housing rate (pp)	-0.340***	-0.282**	-0.087
Secondary housing rate (pp)	0.050	0.105	-0.018
Furnished housing rate (pp)	-0.001	0.035	-0.043*
Selling rate (pp)	-0.067	-0.020	-0.045
Number of blocks	26,269	5,501	7,309
Number of housing units (2000)	5,896,135	1,439,384	2,533,624

Source : Filocom, Perval. Author's calculations.

Sample : treated and control blocks (all types of housing units).

Note : the results are obtained through propensity score kernel matching with optimal bandwidth. The propensity score is estimated separately for each column.

Reading : *** : significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.

Table A11: Heterogeneity with respect to the initial level of price – household characteristics

	All	Low initial price	High initial price
Household size (consumption units)	-0.003**	0.000	-0.003
Nb of children < 18 y.o. per household	0.000	-0.001	0.001
Nb of children < 6 y.o. per household	-0.002	-0.001	-0.002
Household earnings per consumption unit (€)	-235*	-432*	-224
Owner-occupied (pp)	0.122	0.044	0.175
Rented (pp)	-0.067	-0.144	-0.062
Rented social housing (pp)	0.016	0.106	-0.002
Number of blocks	26,269		
Number of housing units (2000)	5,080,150		

Source : Filocom, Perval. Author's calculations.

Sample : treated and control blocks (principal housing only).

Note : the results are obtained through propensity score kernel matching with optimal bandwidth. The propensity score is estimated separately for each column.

Reading : *** : significant at the 1% level; **: significant at the 5% level; *: significant at the 10% level.