Energy efficiency on the European residential sales market: a Meta-Analysis

Numa Bosc - Deborah Leboullenger

May 2021

Abstract:

Asset-backed loans such as mortgages take a significant part of EU banks portfolio and these assets can be affected by climate physical or transition risks. In existing academic literature, energy performance is a real estate price factor particularly scrutinized as more and more evidence shows that energy efficient houses are valued with a price differential we call a "green premium". Reciprocally, least energy efficient assets can suffer from a "brown discount" or become stranded. The quantification of such differential can help prudential authorities substantiate evidence of a risk differential towards climate-related assets the implementation of fine-tuning measures within the prudential framework if relevant. Meta-analysis can be an interesting tool to review and extract a robust quantitative assessment when existing literature is abundant. To prudently quantify a price differential for the European real estate market we built a metabase of existing academic literature and private studies estimating green premium using hedonic prices models. Our meta-analysis concludes that there is an 8.67% green premium in Europe on the residential sales market and a very high degree of heterogeneity. We also found that Energy Performance Certificates (EPC) are the most widely used energy performance labels in EU markets and yields higher premiums. It was also found that peer reviewed journals tend to publish more conservative estimates.

Contents

1 Introduction: Motivations and Context

Environmental, Social and Governance issues are growing in importance within the financial and banking system. Prudential authorities are investigating how climate relevant risks are correctly integrated and whether and how to refine the existing prudential framework to ensure financial stability in the new normal sustainability and climate risks nexus. The evidence of a risk differential between assets that are environmentally aligned and those that are neutral or even harmful can ground such a need for prudential framework adjustments.

The real-estate sector is exposed to climate related risks via two channels:

- Physical risks: Properties could be exposed to climate change events whether they are acute (extreme natural catastrophes increasing in number and intensity with global warming, such as flooding events) or chronic (increasing sea levels and deteriorating land output and productivity). Acute or chronic physical risk can impact house prices, even in neighbourhoods of the event occurring, and borrowers' ability to service their loan payments (due to insufficient protection against natural disasters, for example). Banks are exposed to physical risks as much as the loan's collateral or the borrower's income is exposed to a value deterioration.

- Transition risks: this second type of risk derives from political intervention aimed at curbing the effects of climate change. Real-estate for instance, is targeted by environmental regulation as energy-inefficient dwellings could be forbidden from entering the rental or sales markets. Those dwellings therefore turn stranded meaning temporarily out of the market. Such measures could represent a risk as the inability to sell or rent an energy-inefficient dwelling could prevent a borrower from honouring a loan contract.

While literature remains scarce on the effect of energy performance on probability of default for real estate loans, there is a profusion of literature on the impact of energy performance on property prices .Property prices have an effect on a loan risk profile as it is the denominator of the loan to value (LTV) ratio and LTV buckets are used to classify loans in the Standardised approach for risk weighing in CRR. A property premium or down value given its energy performance in turn impacts the risk profile of a mortgage and eventually the level of prudential capital that a bank can allocate for this loan and asset. This paper aims at quantifying for the European market the level of premium/down value that is to be expected with respect to the energy performance in real estate (from very low performance to energy high efficiency).

2 Methodology

2.1 Presentation of the Meta-Analysis

We conduct a meta-analysis on the existing literature to assess using an econometric method the most reliable level of premium to be expected. This work is based on two meta-analysis studies performed on the same topic from Cespedes-Lopes et al. (2019) and Fizaine, Voye et Baumont (2018). The analysis aims to also keep in mind any potential differences that may exist between ‘industry and ‘academic/public’ studies (the former tend to focus on the ‘positive’ risk differential between ‘green’ or ‘dark green’ vs neutral rather than a ‘negative’ risk differential between ‘brown’ and neutral).

The term ”meta-analysis” was coined for the first time by Gene V. Glass in Primary, Secondary, and Meta-Analysis of Research , an article published in 1976 by the American Educational Research Association. Focusing his needs on education, and confronted with greatly varying results, Gene V. Glass advocated the necessity to extract knowledge by producing “a statistical analysis of a large collection of analysis results from individual studies” . This desire is primarily linked to a paradox of the information age. Indeed, when too many scientific outcomes are to be considered, their heterogeneity will probably increase uncertainty towards knowledge rather than narrowing it. Therefore, since Glass (1976) and regardless of the scientific discipline, an important number of meta-analyses has been carried out and it “has become the practice for evaluating the current flood of conflicting scientific evidence” according to Stanley (2001) .

2.2 Building the meta-base

Building the meta-base slightly differs from a traditional literature review. Performing a meta-analysis on body of literature first require to for that body of literature. Studies may not be selected arbitrarily and details as to how they were retrieved must be clearly enunciated. To that end, Kahn *et al.* (2003) develop a five-step procedure to achieve what they call “systematic reviews”: Framing the question, Identifying relevant publications, Assessing study quality, Summarizing the evidence, Interpreting the findings.

2.3 Selection of new articles

The first step in the process of building this meta-base consisted in merging the meta-bases from the two published meta-analyses mentioned above. Additions to the meta-base were made following the available guidelines to build a systematic review (Kahn *et al.*,2003 or Kahn *et al.*, 2019, Havranek *et al.* 2020) but also from common practices identified from meta-analyses in economics. Indeed, more recent works were found in the *Google Scholar* database, using the following keywords: ”EPC green premium” or ”Green label housing” and in the *Science Direct* database using the keywords "EPC Hedonic premium" and "Energy efficiency premium". Additional estimates were also found using the references section in primary studies already included in the meta-base.

There are two key measures to retrieve from every publication in order to build an effect size: estimates and standard errors. To retrieve estimates, we needed all the estimates produced from the log-level models of the form:

$$log(p) = \alpha + \beta\mathbf{E} + \delta\mathbf{X} + e \tag{1}$$

where p is the price, E is a measure of energy efficiency and X a vector of control variables as suggested by the Hedonic methodology, Rosen (1974). This log-level model will allow one to interpret the estimate as a $100.\beta$ percent increase in price for a unitary increase in the energy efficiency measure.

Commonly, authors propose several estimates as they test various specifications. In that case, we select estimates that are controlling for the most parameters. Whenever the sample is split and the same model is run on sub-samples distinguishing varying Hedonic housing features, all the estimates were collected as long as they are statistically significant and were included in the metabase. However, to prevent double-counting, some were carefully excluded from the final sample.

The other key measure that needs to be retrieved in order to carry out a meta-analysis is the standard error. The availability of such a measure is crucial as the entire meta-analysis is built on an inverse variance modeling technique. All the primary studies that did not disclose it could not be used because their effect size could not be weighted. They were therefore discarded from the metabase unless they gave the *t-statistic* or a confidence interval instead. Retrieving the standard error from an estimate and its *t-statistic* is fairly easy as the two quantities are related as follows:

$$SE = \frac{\beta}{t} \tag{2}$$

When standard errors are not readily retrievable from publication (as it is the case for the French study published by the Notaires de France), we compute "derived standard errors" using the available confidence intervals of the estimated premia from that study [1].

3 The final sample and descriptive statistics

This meta-analysis investigates more closely the European sales market but the initial meta-base compiles 273 effect sizes in the whole world. This number fell to 167 estimates in the European region, to 156 net of the unusable values for the premium and to 112 net

[1] The methodology can be found there: https://handbook-5-1.cochrane.org/

of the studies that did not report the standard errors or anything relevant to derive it (a t-statistic for example). Only 8 studies focused on commercial buildings thus leaving us with 104 residential estimates of a green premium in Europe. Sub-sampling from the complete cases to distinguish between market types yielded 84 data points for the sale contracts and only 26 for the rental market. Summing those two latter values does not add up to 104 because 4 estimates did not specify the contract type and were considered to focus on "both" types of contracts. They were discarded to avoid including an unrelated effect size. One issue must be addressed: in the 84 data points, some of them were counted multiple times. Indeed, as most of the effect sizes were reported for each study, the meta-base included full sample estimates and sub-sample estimates. For instance Fuerst *et al.* 2013 reported the full sample estimate in the United Kingdom but also estimates for specific types of dwellings or specific areas. For that very study, the full sample estimate (the whole United Kindom dataset) did not make it to the final sample. Regional values were reported instead as all the other studies focused on urban areas (Barcelona or Turin for instance). Once all the inter-related values were removed, our final meta-base gathered 69 effect sizes from 19 studies focusing on residential sales on the real estate markets.

Table 1: Sample size

Categories	N
Europe	167
Complete cases	112
Residential obs.	104
Commercial obs.	8
Sale contracts	84
Net of double counting	69
Rental contracts	26

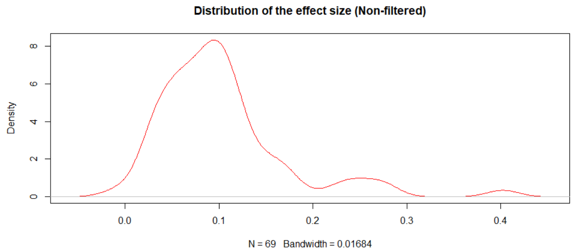
The table in Appendix presents the 15 countries that are present in the meta-base. The second column indicates the mean effect size per country and the third column indicates the number of estimations that have been gathered per country in the primary studies. We can see that two countries are overly represented in the meta-base: France and the United Kingdom with 23 and 22 effect size estimates respectively.

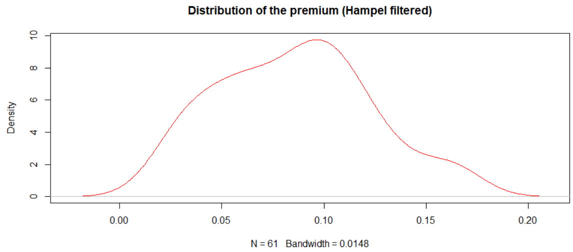
3.1 Outliers Detection

No further analysis should be performed before the normality of the European premia is confirmed. Two issues needed to be tackled: outliers are present in the effect sizes and it is unlikely that said premia are normally distributed. This intuition is confirmed by the results of two Shapiro-Wilk tests that both reject the null hypothesis at the 1 % level.

A Hampel filter procedure was implemented to remove outliers and a Shapiro-Wilk test was performed again which, this time, failed to reject the null hypothesis with a p-value = 0.16 and N=61.

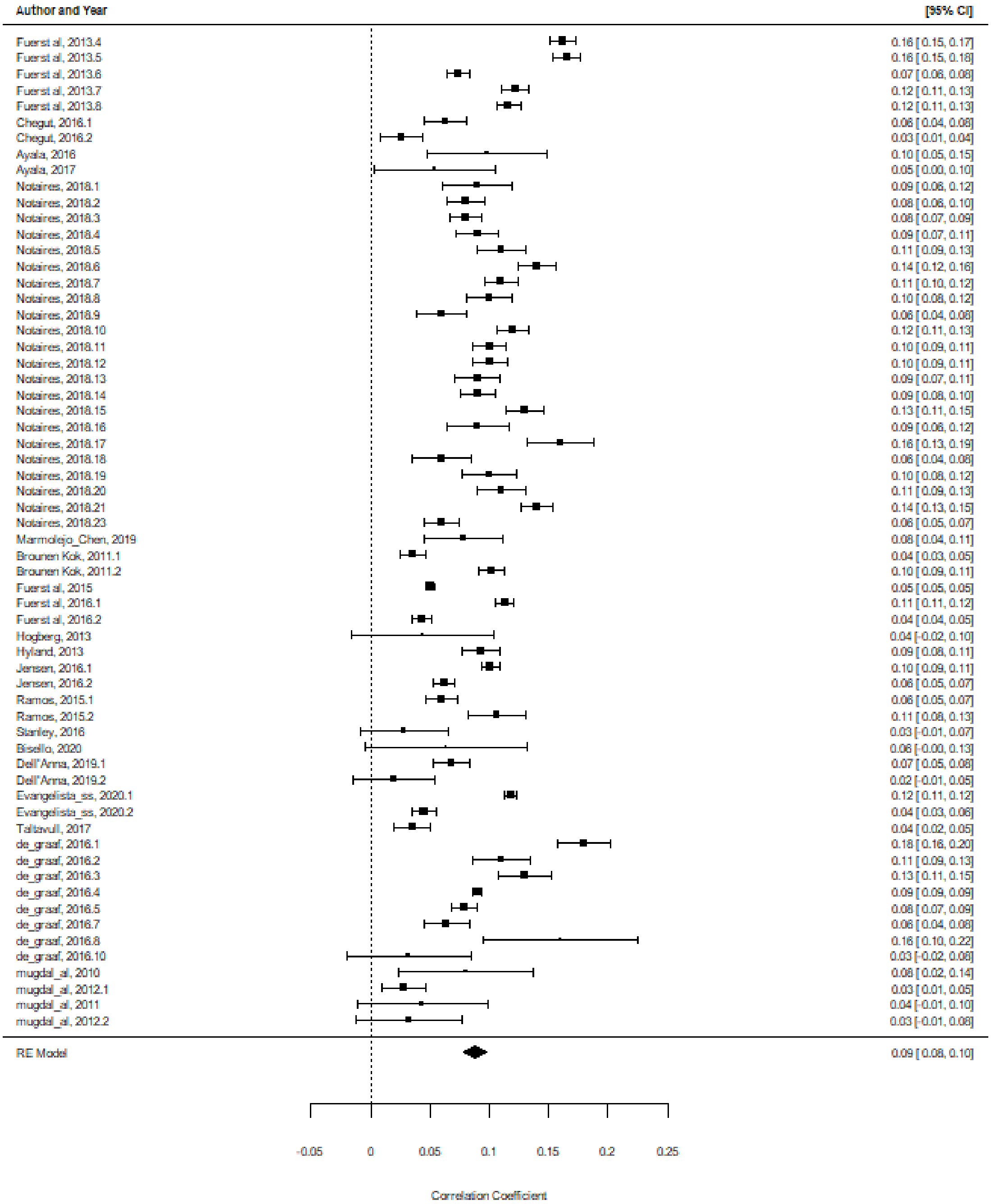
Overall, three procedures were implemented and the Hampel filter one was retained as it provided allowed to obtain a narrower distribution of the premium. The following meta-analysis will therefore be based on the Hampel-filtered metabase for the sales contracts sub-sample. The density plots available below illustrate the impact of the Hampel filter procedure on the distribution of the effect sizes.





4 Meta-Analysis results

The sample was characterized with a very high level of heterogeneity, which is not uncommon for a meta-analysis focusing on the green premium. The random effect modeling technique will therefore be the one considered, supported by an investigation of such heterogeneity.



Indeed, the $I^2 = 97.19$ so it is safe to say that sampling error has an almost null influence on the variation in the premia. This differs from some of the studies from other fields reviewed in the first part mainly because econometrics techniques are oftentimes not performed in a controlled laboratory. Therefore the conditions of a study may never be exactly reproduced. This degree of heterogeneity is consistent with that in Cespedes-Lopes *et al.* 2019. Overall, given the high level of true-heterogeneity, one will then have to turn to methodological strategies used in the primary studies in order to further comprehend such discrepancies. The overall estimated premium for the sales market in Europe, measure through EPC labels within hedonic regression models is equal to 8.79% within the following confidence interval: $[0.0784; 0.0974]$ with a p-value below 0.0001. This is given the bibliography that was gathered and the available data it provided. It is obviously subject to change. But it seems that the studies contained in our systematic review confirm the existence of an 8.79% green premium at the scale of Europe.

Sub-sampling in Europe and then sub-sampling for a specific market type or contract type, did not cause our data to become more homogeneous. In the case of the green premium, sub-sampling at the regional or at the country level can make sense. That way, a common underlying effect may become more and more visible in the data. Such an underlying effect could be the energy efficiency legislation that is likely to be the same at the country level.

As the rule of thumb to carry out a meta-analysis requires a minimum of 10 effect sizes, it is possible to carry out meta analyses on sub-samples to assess heterogeneity. The rationale behind this is see if this could lead to a lower heterogeneity level. Since 23 effect sizes were available for France, another meta-analysis was performed for this country. It was found that the green premium are higher in France than at the European level. Indeed the premium derived from the sub-sample is equal to 9.78 within the following confidence interval: [8.82;10.74]. The standard error however, is almost the same between the two meta-analyses: 0.0049 in France and 0.0048 for the full sample.

The first point to be tackled is the fact that almost all of these estimates share common authorship as the study is carried out every year by a consultancy based on notarial data. Hence we observe very little between-study variation and a more homogeneous set of estimates. Indeed, sub-sampling at the country level for France lowered the I^2 level to 89.79%. However the heterogeneity level is still high. The regression and tests performed on the French sub-sample is to illustrate all the theoretical underpinnings related to methodological heterogeneity. Put differently, it is very likely that if additional primary studies were available for the French markets, it would increase the degree of heterogeneity. Indeed the estimation methodologies would certainly vary more. [2]

A single country or regional analysis also allows to investigate the geographical heterogeneity factors. Climate, building stock and construction norms tend to be more homogeneous within countries or even regions. If a certain category of buildings is targeted

[2]The related forest plot can be found in the Appendix

by a public policy through construction norms, changes in energy efficiency are observed at country level with the construction year being the heterogeneity factor rather that the geographical location. The identification of those factors is a first step to determine if a price premium or discount can be applied to all EU countries in a harmonised way from Lisbon to Stockholm.

5 Meta-regression

5.1 Model specification

Meta regression is a technique that is commonly used to investigate heterogeneity following the results given by a meta-analysis. It requires to use the usual standard linear model:

$$\mathbf{E}[Y|X] = X.\beta \tag{3}$$

Based on the formulation of research hypotheses regarding potential drivers of our effect-sizes variation, relevant metadata is extracted from the primary studies. As for a classic linear regression, we want to collect data for our variable of interest while controlling for other characteristics. Equation (8) can therefore be written:

$$\mathbf{E}[Y|X, Z] = X.\beta + Z.\gamma, \tag{4}$$

$$Y \sim \mathcal{N}(\mu, \sigma^2) \tag{5}$$

Where X is a vector of variables of interests and Z a vector of control variables. Based on Cespedes-Lopez *et al.* (2019) the vector of control variables for each paper contains the following data:

- non labled comparison (dummy): codes 1 if the premium is estimated by comparing EPC labeled dwellings to non labeled dwellings. Codes 0 otherwise.
- dwelling (dummy): dwe_i codes 1 if the regression in study i controls for the usual hedonic criteria of the real estate goods studied. Codes 0 if no hedonic criteria are controlled for.
- building (dummy): bui_i codes 1 if the regression in study i controls for the external amenities of the real estate goods considered. Codes 0 otherwise.
- neighborhood (dummy): nei_i codes 1 if the regression in study i controls for neighborhood characteristics such as social and economic status of the neighborhood the dwellings are part of. Codes 0 otherwise.
- location (dummy): loc_i codes 1 if the study i considers the exact location of the dwelling. Codes 0 otherwise.

- market (dummy): mar_i codes 1 if the study i controls for market characteristics of the dwelling such as its time spent on the market for instance.
- construction date (dummy): codes 1 if the study i controls for the age of the dwelling. Codes 0 otherwise.
- multifamily (dummy): codes 1 if the study i focuses on multi-family buildings and codes 0 if it focuses on single-family buildings (apartment or houses).
- number of authors (discrete): indicates the number of authors included involved in the primary study.

Introducing a vector of controls, also called moderators in the context of a meta-analysis allows one to focus on the different aspects of a methodology. Indeed, with such a high degree of heterogeneity being assigned to methodological discrepancies between primary studies, it is necessary to account for such differences. For instance, the study by Copenhagen Economics (2015) states that *"controlling for reports on errors and omission as well as the building year is very important and it decreases the effect of energy standards significantly"*. One may therefore expect more conservative results from studies that are able and willing to introduce a higher number of control variables in the hedonic model they regress.

5.2 Research Hypotheses

Based on the level of heterogeneity reported in the meta-analysis, it appears that variability in effect sizes is driven by the variability in statistical methodology between the various primary studies. And since all the studies we gathered deal with Energy Performance Certificates (EPC), we hypothesize that different treatments of variable E in Equation (6) will yield varying levels of premium. Simply looking at all the regression tables gives insight about the way researcher explored the explanatory potential of energy efficiency on transaction prices. There are two types of studies: those that work with sub-groups of EPC categories and those that do not. In fact, one may expect varying levels of significance, and thus varying impacts of EPC labels on transaction prices when using a group variable made of classes A, B and C versus the lower ones than using six separate classes.

Variables of interests were defined and collected in order to investigate the varying treatment of EPC as an independent variable in the primary studies.

With regards to grouping and because all the EPC labels considered in primary studies had almost all the same number of EPC categories, we defined a "distance to reference variable". This covariate is expected to be characterized with a positive coefficient. Indeed one expects a larger premium when A-rated dwelling prices are compared to G-rated dwelling prices rather than D-rated dwelling prices. Thus, this variable varies from 1 to 5. For instance, a value of 2 means the premium reported is two steps above the reference category. Formally:

Hypothesis 1: *Primary studies with a higher "distance to EPC reference" are expected to report a higher premium.*

Another aspect of the EPC treatment that is worth investigating is the grouping of label categories. Indeed all the primary studies do not report premia based on a single category. More precisely, the regressions sometimes report a green premium on dwellings labeled A, B or C, others only group A and B or consider each level individually. We expect this variable to be negatively related to the premium since grouping further and further inevitably reduces the distance to reference defined above. Yet, it might be necessary to test the relationships in separate regressions as the two variables are likely to be correlated. The negative sign could be understood as follows: comparing A-rated dwellings to D-rated dwellings (thus group = 1) could yield a higher premium as the average energy efficiency is higher for the A group than for the A-B group.

Hypothesis 2: *There is a significant negative relationship between the level of grouping and the premium, i.e the higher the number of grouped categories, the lower the reported premium for the group*

One key feature of the analyzed sub-sample is that it collects effect sizes from both published works and unpublished works. Indeed, Fuerst *et al.* 2013 is a report requested by the Department of Energy and Climate Change and *La valeur verte des logements* is a study conducted by a private company - Min.not - for *Notaires de France*. It represents. And Ramos 2015 is a working paper. Overall it represents a total of more than 30 effect sizes in the meta-base.

Hypothesis 3: *Given that the conditions for publishing differ between a peer-reviewed scientific journal and a company's research department, estimates reported in scientific papers might be of different magnitude than that of estimates from non-scientific papers*

As suggested we will propose two models to test these hypotheses related to the treatment of EPC rankings in primary studies.

H1:

$$premium_i = \alpha + \beta_1.distrat_i + \epsilon_i \tag{6}$$

H2:

$$premium_i = \alpha + \beta_1.group_i + \eta_i \tag{7}$$

H3:

$$premium_i = \alpha + \beta_1.publi + \gamma_i \tag{8}$$

Where textitdistrat stands for *distance rating* and measures the distance to the reference EPC rank. And finally $group_i$ is the number of EPC categories considered for the reported premium in primary study i. ϵ, η and γ are the three respective error terms.

The final vector of control variables is defined as follows:

$$Z = \begin{bmatrix} non-labeled-comparison \\ building \\ neighborhood \\ location \\ market \\ constructiondate \\ published \end{bmatrix}$$

Indeed some of the control variables mentioned above were dropped due to their lack of variability. For instance, all of the effect sizes were estimated while controlling for dwelling hedonic properties (as required).

The final models used to test the three hypotheses are defined as follows:

$$premium = \alpha_1 + distrat'.\beta_1 + Z'.\delta_1 + e_1 \tag{9}$$

$$premium = \alpha_2 + group'.\beta_2 + Z'.\delta_2 + e_2 \tag{10}$$

$$premium = \alpha_3 + published'\beta_3 + Z'.\delta_3 + e_3 \tag{11}$$

5.3 Results

The table provided in Appendix summarizes the results from the models presented above. Column (1) (2) and (3) correspond to equations (11) (12) and (13). Control variables are introduced from column (4) to column (7) in order to test the robustness of the first estimation with the previously cited moderators. Column (7) only excludes the variable reflecting the group size as it will prove to be less relevant than the "distance to reference" variable to tackle the EPC methodological treatment.

Two estimations were produced. First, the meta-regression was performed using the whole set of control variables. Then, the same meta regression was performed but this time, excluding the variables of interest. Control variables were then selected using the F-test, looking for the maximum number of controls to include while keeping the p-value of the F-test below 5%. Those that remained were thus: "non labeled comparison", "neighborhood controls" and "market controls". Both tables are provided in the appendix. The latter one will be discussed here (Figure 4. in the Appendix).

The table shows tests for the three hypothesis first treated individually. This is only to get an idea of what can be expected. It is also useful because it gives insights to compare Hypothesis 1 and 2. The two variables were also compared using the full set of control variables which confirmed the poor influence of the "group" variable on the magnitude of the premia.

However, one needs to recognize that the explanatory power of the overall meta-regression model is not very high. This may be due to the fact that solely focusing on metadata of the primary-studies is unlikely to account for the whole premia variations. Therefore the results presented here should be understood as a baseline intuition at best. Even though, the meta-analysis concluded that there is indeed a sales premium in Europe, the research factors that could drive the results are far from being investigated in their entirety here, for lack of observable data.

Focusing on the first column, the rationale behind the first hypothesis seems to be valid. The simple regression model indicates a positive and statistically significant relationship between the distance to reference in EPC. Since this estimate is not controlled for, the magnitude and the level of significance should be considered with caution. The results related to the first hypothesis seem rather robust as the variable "distance to reference" remains statistically significant in the three models when controlling for additional study characteristics. This result remains statistically robust when control variables are added: it's coefficient remains statistically significant at the 5% level even when control variables are added.

Regarding the second column, it appears that the size of the group has no significant effect on the premium and if it were to be, there would be a positive relationship between the two quantities. This should be discussed here a little more as grouping may affect the significance level of the premium's coefficient rather than its magnitude. Introducing the moderator variables makes the coefficient of the group size variable even less significant. The variable "distance to reference" will therefore be considered as the variable that best reflects the variations in use of the EPC rankings in the primary studies. It seems that the variable "distance to reference" is more suited than the "group" variable to investigate the EPC related variations of premia.

The third column displays the regression result for the model that tests the third hypothesis. Results published in scientific journals tend to report lower premia. This result is robust as it remains significant at the 1% level when controlling with $\mathbf{Z}$ and at the 5% level when controlling for both the "distance to reference variable" and $\mathbf{Z}$.

The last column reports the result for the complete meta-regression. It's explanatory power is around 22% and it confirms the first and third hypotheses (both coefficients are statistically significant at the 5% level), given the set of moderators introduced. Indeed, increasing the distance between the reference EPC band and the reported band for the premium will lead to higher estimates by approximately 1%, and keeping in mind that the standard error is equal to 0.0049. Results published in scientific journals tend to be lower than those published in non-scientific journals. In other words, research papers report more conservative results regarding the green premium.

Additionally, it appeared relevant to assess the impact of the interaction between the variable "distance to reference" and the variable "published". Different model specifications are available in Appendix. It did provide another result, but which needs to be considered cautiously. The interpretation of an interaction term is slightly different from that of an individual independent variable. When modeling $premium_i = \alpha + \beta_1.dist_i + \beta_2.pub_i + \beta_3.dist_i.pub_i + \gamma.\mathbf{Z} + \epsilon_i$, the effect of "distance to reference" on $premium$ will be $\beta_1 + \beta_3$ whenever the $pub = 1$ and to β_1 whenever $pub = 0$. Also, γ is a vector of coefficients and $\mathbf{Z}$ is the vector of moderators. Here is the result observed from this model: the interaction term between "distance to reference" and "publication" was statistically significant at the 1% level. In addition, the intercept is now statistically significant at the 5% level only, the distant to reference is significant at the 1% level but the publication variables no longer shows significance. So the fact that "distance" is significant indicates that the effect of distance gains in significance whenever $pub = 0$ and assessing that publication is no longer significant indicates that it does not impact the premium when $distance = 0$ which is rather obvious as $distance = 0$ makes no sense considering what it measures - the minimum value this variable can take is 1 by definition. A high p-value (i.e. non significant) does not mean that there is no effect of the publication variable but rather that the sample size is too small. However, it is worth noting that the individual effect of the publication variable is now a positive one, rather than a negative one but it is no longer significant. The impact of "distance to reference" will therefore be equal to $\beta_{dist} + \beta_{publi*dist} = 0.0029$, significant at the 5% level. The overall explanatory power of the model including the interaction term is 27.21%.

The last step in this meta regression is to confirm or overturn those regression results. Indeed a robustness test must be performed on the model as multiple estimates are drawn from the same studies. This is not exactly comparable to the issue mentioned above related to the double counting that could occur if both full samples and subsamples estimates were pooled in the meta analysis. This aims at taking into account the fact that some results come from the same study. This leads to nuance the results developed above.

It appears that the coefficients are not as significant as in the previous regression. Indeed, the coefficients for the variables of interest "distance to reference" and "published" are only significant at the 10% level. Their magnitude however, remains the same. Surprisingly, the variables that indicates whether a study compares top labeled dwellings to non labeled dwellings is now significant at the 5% level. Comparing a top labeled dwelling to a non labeled dwelling implies that the energy efficiency of the latter is not included

Table 2: Robust meta-regression

Variables	Estimates	se	pval
Intercept	0.07165	0.01507	0.000596
Distance	0.01231	0.00568	0.053034
Published	-0.02101	0.01005	0.060551
Non labeled	-0.0406	0.01462	0.017999
Neighbourhood controls	-0.01338	0.01162	0.274141
Market controls	0.00215	0.01053	0.842029

in its price, leading to a degree of overvaluation of energy inefficient dwellings. All else equal, this results in lower estimated premia.

There is still a positive relationship between the distance to reference and the magnitude of the observed premium and a tendency to observe more conservative results in published scientific articles.

6 Discussion

As mentioned in the Results section, the explanatory power of controlled variables of interest in the meta-regression model is just above 25%. Coupling this measure with the level of heterogeneity presented in the forest plot, such results should be considered with caution. It means that this model is only explaining a small share of the premia variation in Europe. A high degree of heterogeneity was expected and observed on a single continent and our meta-regression already explained less than a third of total variation. Going further, more moderator variables can be integrated to account for potential continental or methodological discrepancies.

Another important meta-analysis result is that there is a significant negative relationship between the level of grouping (ie. the granularity of the dependant variable and the control group) and the premium. The meta analysis shows that for any study in the metabase, the higher the number of grouped categories in a study (eg. looking at a green premium on dwellings labelled A, B or C, versus others, only grouping A and B or considering each level individually), the lower the reported premium for the group. This can be expected for two reasons: one is mechanical as grouping inevitably further reduces the distance to reference, the second is due to the heterogeneity within groups including several EPC scores (C rated buildings diluting the premium effect of A-rated building, having much lower energy performance standards). Nonetheless, due to a lack of data available, for numerous studies, very few observation are found using a granular classification especially at high efficiency scores (A or B) and letter by letter analysis has proven challenging because it faced significance issues. This result is important from a policy perspective for two reasons. The first is that, for the existing stock of historical data on EPCs, it has been challenging to find evidence of a granular (i.e. letter by letter) evidence of a valuation differential due to the lack of (quality) data availability and coverage of the existing building stock especially at high energy efficiency levels. Second, we see that grouping several EPC categories (e.g. A, B and C) to D-rated dwellings can lead to a dilution in the premium relative to comparing A-rated dwellings to D-rated dwellings (thus group = 1) that yield a higher premium, as the average energy efficiency is higher for the A group than for the A-B-C group. Going further, as more data becomes available, studies will gain insight in testing this relationship in separate regressions in a letter by letter correlation analysis to refine and better quantify the difference in asset values for each EPC score.

Last but not least we acknowledge in this analysis two shortcomings that should be further investigated with more granular and reliable data on energy efficiency and property valuation. First, the metabase does not contain enough estimates for property discounts (or down-values) for the least energy performing segment (F and G rated properties). Second, not all EU countries are represented in the sample and there is a potential bias towards regions and markets that provided the most estimates (in the case at hand, the real estate market in France represents 50% of the metabase). Some data points in Belgium were discarded because they were based on a measure that is different from EPC. Indeed, former meta-analyses pooled estimates from all the continents, regardless of the market type (residential or commercial) but more importantly regardless of the label. Including various measures of energy efficiency captured by different labels is likely to max out the level of heterogeneity and make the comparison between studies more difficult.

7 Conclusion

The meta-analysis and the meta-regression performed in this paper supports the existence of a premium linked to energy efficiency as measured by the EPC labels. The high degree of heterogeneity led to an investigation of the premia variation that turned out to be significantly influenced by two things. The first one is the tendency of academic journals to publish more conservative results and the second one reflects methodological choices regarding the EPC reference bands. It is clear that additional meta-data should be introduced in order either to exploit new research hypotheses or control the latter. Indeed, a great share of the premia variation explanatory variables remains unobserved.

Moreover the role of this type of analysis is to produce knowledge in order to assess the relevance a banking-fine tuning regarding capital requirements.Meta analysis is an interesting tool when assessing a vast existing literature and to extract a prudent quantitative assessment that can be used as a benchmark at EU level while being aware of the quantification caveats in place in the market. This could encourage researchers and institutions to carry out further econometric analysis on existing quantitative literature such with the use of panel data for example. Feeding the metabase with most recent research can also help to assess the effect of time, data coverage and quality on the dependent variable in scrutiny.

8 Appendix

8.1 Descriptive table

	Country	Mean Effect Size (%)	N
1	austria	6.200	2
2	belgium	4.418	11
3	denmark	6.325	4
4	france	10.565	23
5	germany	6.950	2
6	ireland	3.183	6
7	italy	6.575	2
8	netherlands	5.300	9
9	norway	6.617	11
10	portugal	11.682	9
11	romania	3.500	1
12	spain	3.287	6
13	sweden	5.205	2
14	switzerland	18.290	2
15	uk	11.869	22

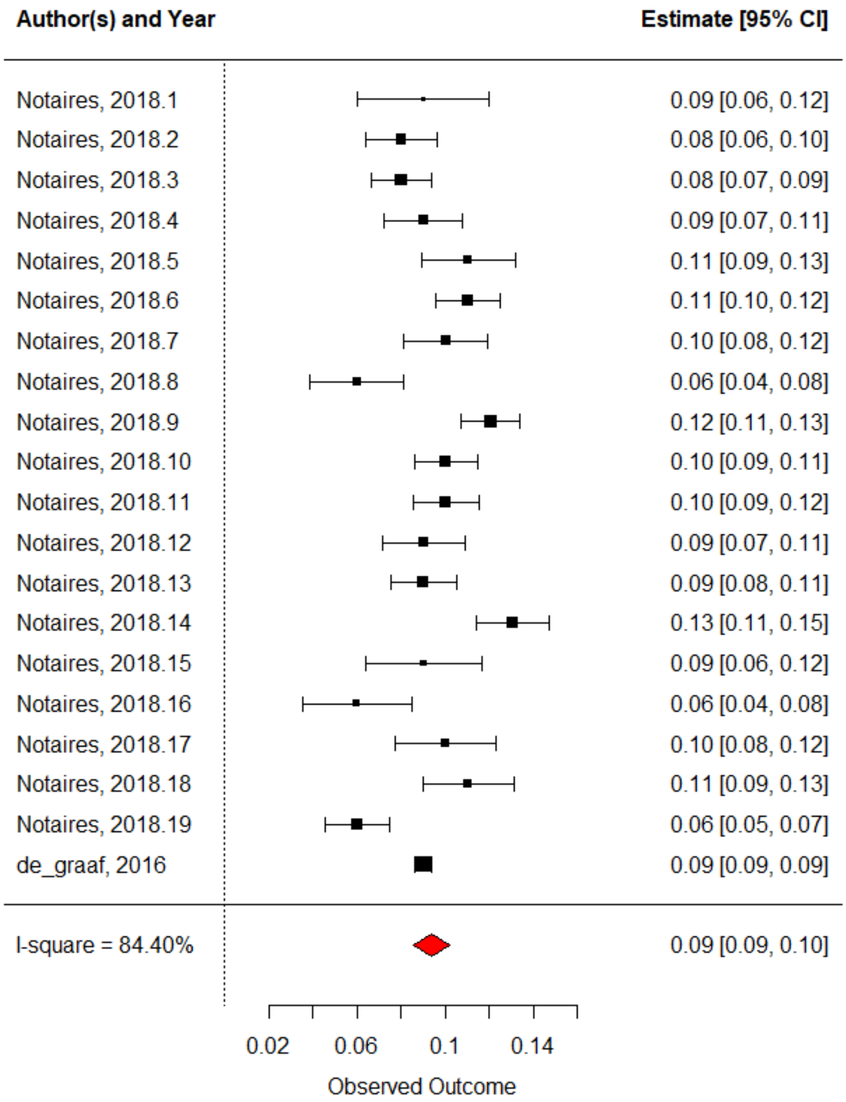
Table 3: Northern America

Country	Mean Effect Size (%)	N
Canada	12.200	1
USA	5.233	52

Table 4: Asia

Country	Mean Effect Size (%)	N
China	−0.220	2
Japan	−5.773	3
Singapore	7.480	3

8.2 Forest plots



8.3 Meta-regression results

	1	2	3	4	5	6	7
Intercept	**0.0632**	**0.0717**	**0.0989**	**0.0487**	**0.0787**	**0.1142**	**0.0776**
(s-e)	*0.011*	*0.0133*	0.0059	*0.0204*	*0.0155*	*0.0171*	*0.0243*
(p-val)	*0.0001*	*0.0001*	0.0001	*0.0203*	*0.0001*	*0.0001*	*0.0023*
Distance	**0.001**	-	-	**0.0126**	-	-	**0.0112**
(s-e)	*0.004*	-	-	*0.0055*	-	-	*0.0054*
(p-val)	*0.0129*	-	-	*0.0266*	-	-	*0.0429*
Group	-	**0.0085**	-	-	**0.0064**	-	-
(s-e)	-	*0.007*	-	-	*0.0077*	-	-
(p-val)	-	*0.2251*	-	-	*0.4103*	-	-
Published	-	-	**-0.0271**	-	-	**-0.0309**	**-0.0276**
(s-e)	-	-	*0.0088*	-	-	*0.0136*	*0.0133*
(p-val)	-	-	*0.0021*	-	-	*0.0267*	*0.0424*
Non labeled	-	-	-	**-0.0559**	**-0.079**	**-0.0496**	**-0.0294**
(s-e)	-	-	-	*0.0321*	0.0315	*0.0332*	*0.0336*
(p-val)	-	-	-	*0.0876*	0.0151	*0.1409*	*0.3853*
Building controls	-	-	-	**0.0009**	**-0.0086**	**-0.0263**	**-0.0169**
(s-e)	-	-	-	*0.0112*	*0.0111*	*0.0135*	*0.0138*
(p-val)	-	-	-	*0.9377*	*0.441*	*0.0562*	*0.2261*
Neighborhood controls	-	-	-	**-0.0187**	**-0.0155**	**-0.0073**	**-0.0113**
(s-e)	-	-	-	*0.0135*	*0.014*	*0.0139*	*0.0136*
(p-val)	-	-	-	*0.1724*	*0.2735*	*0.5992*	*0.4089*
Location controls	-	-	-	**-0.0092**	**-0.0142**	**-0.0087**	**-0.0074**
(s-e)	-	-	-	*0.0119*	*0.0129*	*0.0119*	*0.0115*
(p-val)	-	-	-	*0.4399*	*0.2782*	*0.4684*	*0.5211*
Market controls	-	-	-	**0.0089**	**0.0296**	**0.0111**	**-0.0045**
(s-e)	-	-	-	*0.0193*	*0.0182*	*0.0189*	*0.0198*
(p-val)	-	-	-	*0.6473*	*0.1089*	*0.5618*	*0.8213*
Construction date	-	-	-	**0.0188**	**0.0101**	**0.0075**	**0.0135**
(s-e)	-	-	-	*0.0134*	*0.014*	*0.0133*	*0.0133*
(p-val)	-	-	-	*0.1668*	*0.4717*	*0.5768*	*0.3135*

Figure 1: Meta-Regression results

	Interact
Intercept	0.0454
(s-e)	0.029
(p-val)	0.0123
Distance	0.0254
(s-e)	0.0091
(p-val)	0.0071
Group	-
(s-e)	-
(p-val)	-
Published	0.0411
(s-e)	0.0379
(p-val)	0.2823
Non labeled	-0.0334
(s-e)	0.0328
(p-val)	0.3134
Building controls	-0.0043
(s-e)	0.015
(p-val)	0.775
Neighborhood controls	-0.02
(s-e)	0.014
(p-val)	0.1581
Location controls	-0.016
(s-e)	0.0121
(p-val)	0.1907
Market controls	-0.0134
(s-e)	0.0198
(p-val)	0.5039
Construction date	0.0067
(s-e)	0.0134
(p-val)	0.6207
Distance . Pub=1	-0.0239
(s-e)	0.0124
(p-val)	0.059

Figure 2: Interaction: Distance - Publication

Bibliography A: Statistical methodology

- Glass, Gene V. Primary, Secondary, and Meta-Analysis of Research . Educational Researcher, vol. 5, no 10, November 1976, p. 3 8.
- Stanley, T. D. Wheat from Chaff: Meta-Analysis as Quantitative Literature Review. The Journal of Economic Perspectives, vol. 15, no 3, 2001, p. 131 50.
- Borenstein, Michael, éditeur. Introduction to Meta-Analysis. John Wiley Sons, 2009.
- Deeks JJ, Higgins JPT, Altman DG (editors). Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane, 2021.
- Frisch, Ragnar. Editor's Note . Econometrica, vol. 1, no 1, 1933, p. 1 4
- Khalid S. Khan, et al. Five Steps to Conducting a Systematic Review. 2003, p. 4.
- Moher, David, et al. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. Journal of Clinical Epidemiology, vol. 62, no 10, octobre 2009
- PRISMA-P Group, et al. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 Statement . Systematic Reviews, vol. 4, no 1, décembre 2015, p. 1. DOI.org (Crossref), doi:10.1186/2046-4053-4-1.
- Havránek, Tomáš, et al. Reporting Guidelines for meta-analysis in economics. Journal of Economic Surveys, vol. 34, n 3, juillet 2020, p. 469-75. DOI.org (Crossref), doi:10.1111/joes.12363.
- Higgins, Julian P. T., et Simon G. Thompson. Quantifying Heterogeneity in a Meta-Analysis . Statistics in Medicine, vol. 21, n 11, juin 2002, p. 1539-58. DOI.org (Crossref), doi:10.1002/sim.1186.
- Cochran, William G. The Combination of Estimates from Different Experiments . Biometrics, vol. 10, n 1, mars 1954, p. 101. DOI.org (Crossref), doi:10.2307/3001666.
- DerSimonian, Rebecca, et Nan Laird. Meta-Analysis in Clinical Trials . Controlled Clinical Trials, vol. 7, n 3, septembre 1986, p. 177-88. DOI.org (Crossref), doi:10.1016/0197-2456(86)90046-2.
- Higgins, J. P. T. Measuring Inconsistency in Meta-Analyses . BMJ, vol. 327, n 7414, septembre 2003, p. 557-60. DOI.org (Crossref), doi:10.1136/bmj.327.7414.557.
- Borenstein, Michael, et al. A Basic Introduction to Fixed-Effect and Random-Effects Models for Meta-Analysis . Research Synthesis Methods, vol. 1, n 2, avril 2010, p. 97-111. DOI.org (Crossref), doi:10.1002/jrsm.12.
- Nikolakopoulou, Adriani, et al. Demystifying Fixed and Random Effects Meta-Analysis . Evidence Based Mental Health, vol. 17, n 2, mai 2014, p. 53-57. DOI.org (Crossref), doi:10.1136/eb-2014-101795.
- Nelson, Jon P., et Peter E. Kennedy. The Use (and Abuse) of Meta-Analysis in Environmental and Natural Resource Economics: An Assessment . Environmental and Resource Economics, vol. 42, n 3, mars 2009, p. 345-77.

- Stanley, T. D., et Stephen B. Jarrell. Meta-Regression Analysis: A Quantitative Method of Literature Surveys: Meta-Regression Analysis. Journal of Economic Surveys, vol. 19, n 3, 1989, p. 299-308.

Bibliography B: Economics meta-analyses

- Johnson, Noel D. Trust Games: A Meta-Analysis. Journal of Economic Psychology, 2011, p. 25.
- Thomas van Huizen , Janneke Plantenga , Do Children Benet from Universal Early Childhood Education and Care? A Meta-Analysis of Evidence from Natural Experiments, Economics of Education Review (2018)
- Alinaghi, Nazila, et W. Robert Reed. Taxes and Economic Growth in OECD Countries: A Meta-Analysis . Public Finance Review, vol. 49, no 1, janvier 2021, p. 3 40. DOI.org (Crossref), doi:10.1177/1091142120961775.
- Doucouliagos, Hristos, et Patrice Laroche. Unions and Profits: A Meta-Regression Analysis 1 . Industrial Relations: A Journal of Economy and Society, vol. 48, n 1, janvier 2009, p. 146-84. DOI.org (Crossref), doi:10.1111/j.1468-232X.2008.00549.x.
- Disdier, Anne-Celia, et Keith Head. The puzzling persistence of the distance effect on bilateral trade. The Review of Economics and Statistics, 2008, p. 12.
- Cespedes-Lopez, et al. Meta-Analysis of Price Premiums in Housing with Energy Performance Certificates (EPC). Sustainability, vol. 11, n 22, novembre 2019, p. 6303. DOI.org (Crossref), doi:10.3390/su11226303.
- Conseil Français de l'Energie, Programme Energie et Territoire, "Habiter: Recherches sur l'évaluation hédonique des interactions habitat-énergie-territoire", *Rapport final*, Décembre 2016.
- Abdullah, Abdul, et al. Does education reduce income inequality? A meta-regression analysis Journal of Economic Surveys, vol. 29, n 2, avril 2015, p. 301-16.
- Daniel, Vanessa E., et al. Flooding Risk and Housing Values: An Economic Assessment of Environmental Hazard . Ecological Economics, vol. 69, n 2, 2009, p. 355-65.