2021/2022

Msc. Econometrics and Management Science, Quantitative Finance track,
Master Thesis

Corporate green bond analysis between the US, China and Europe with the account of time-invariant heterogeneity

Authors:
Lin Huiying

Student Number and track:
618583, QF

August 7, 2022

Supervisor: prof.dr. Michel van der Wel
Second assessor: dr. Wendun Wang

JEL Codes: C58, G12, G24, O16
Key words: Green Bonds, Time-invariant Heterogeneity, the Hausman Taylor Model, pooled OLS regression, Corporate Bonds, Size Premium, GDP, the Garch model

Contents

Abstract

The paper investigates the impact of the green bond label on corporate bonds that are issued between 2019 to 2021. Furthermore, the impact is compared between China, the US and Europe. The analysis further extends to explore the effect of the GDP factor and firm size and the relation between these factors and the green label. On top of that, the paper applies the Hausman Taylor model to account for time-invariant heterogeneity. It is found that pooled OLS outperforms the Hausman Taylor model, thus time-invariant heterogeneity does not disturb the green bond analysis. The empirical results present that green bonds are more rewarding in the US and Europe than conventional bonds but that result holds the opposite in China. The results differ from Zerbib [2019] where the research is established in 2019 and the author finds the negative greenium. Moreover, green bonds in China and Europe are more resilient during an economic low and perform much better than in normal economic situations. Green bonds in the US show no difference under changing in economic circumstances. Lastly, green bonds in Europe that are issued by small firms give an extra yield of 0.0401 bps compared with conventional bonds.

1 Introduction

Green bonds are a popular asset class arising over the most recent years. With the increasing awareness of climate change and environmental protection, countries and corporates are working towards more green economic and business transformations. Green bonds are one big core part of the credit market that imparts the idea of supporting sustainability. Green bonds are issued by corporates or government-backed entities to borrow money from the public and in return give coupons or interest rates in terms. The money pooled via green bonds is used for facilitating improvement in the projects related to the environment and climate. The issue of the very first green bond is retrieved back in 2007, whom the issuer institution being European Investment Bank and World Bank with the goal of supporting the eligible climate-focused projects in Sweden. Ever since the green bonds market started blooming and the volume is increasing magnificently worldwide. In 2013, the first corporate bond was issued by a Swedish real estate company named Vasakronan. The program intends to support projects of lower energy use and lower climate impact. Ever since, meta corporates such as Apple, Unilever and ICBA responded quickly by issuing a great number of green bonds. Green bonds become more preferred over the years for a few main reasons. First, it results in lower borrowing costs. Zerbib [2019] finds a negative greenium which explicitly points out that green bonds give out lower yields compared to conventional bonds. Gianfrate and Peri [2019] also state the findings that issuers such as companies can gather funds for refinancing or investments in a cheaper way. Moreover, companies tend to improve or transform to a more sustainable and environmental-friendly image by issuing green bonds. This is also referred to as 'Green washing'. Tang and Zhang [2020] present the results that the issuing of the first green bond stimulates the increment in the stock price from the same issuer. The issuing of the first green bond sends a signal to the investors that the corporate plans to have future sustainable projects. The investors, therefore, have a stronger faith in the issuer corporate's future. In addition, the existence of green bonds also caters to the appetite of the market where some investors demand more ethical investment options. However, the differences in the green bond label across countries and regions are rarely investigated. Therefore, the underlying attention of this paper is to inspect the direct impact of the green label and compare them across countries. Moreover, the analysis will take economic factors such as GDP and firm size into account.

Figure 1 presents that green bonds are springing out significantly in recent years, especially in the year 2021 with the total corporate green bonds issue amounting that over 350 billion. What stands out as well is that China and the US are the top 2 countries with the highest volume of issuing green bonds from 2019 and onward. Meanwhile, France, Sweden, Germany and Netherlands are also active issuers with a great number of issues. And what is the definition of 'green' and who gets to decide? There are two green bond guidelines are universally adopted and acknowledged by most companies and countries: Green Bond Principles (GBP) and Climate Bonds Initiative (CBI). GBP was initialized by multiple investment banks including Deutsche Bank, Goldman Sachs, HSBC, JPMorgan Chase etc in 2014. Further, International Capital Market Association (ICMA) took over for monitoring and adjustments of the guidelines. 4 core components are covered in the GBP: use of proceeds, the process for project evaluation and selection, management of proceeds and reporting. However, the details of defining 'Green' are not revealed to the public but remained exclusive to issuers. Alternatively, CBI is an investor-focused not-for-profit organization whose standards for green bonds are widely adopted by the majority of countries. CBI offers very specific green taxonomy sectors so the investors or issuers can assess with these standards and confirm with qualifications of 'green bond'. In this paper, data are provided from **Eikon** [1] and the standards of green bonds are in line with GBP.

[1]Eikon is a set of software products provided by Refinitiv for financial professionals to monitor and analyze financial

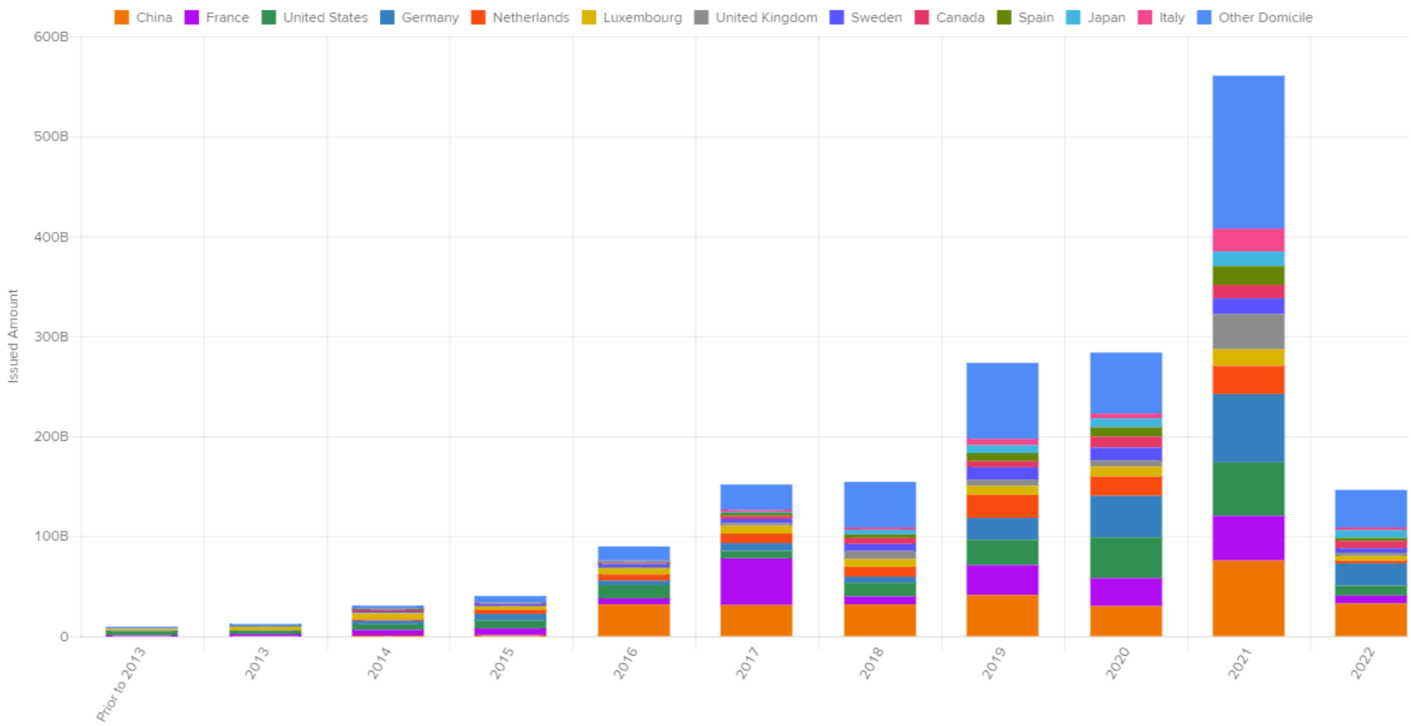


Figure 1: Corporate Green Bonds Overview, source: Eikon Green Bond Guide.

As the demands for green bonds increase rapidly, it is vital for investors to learn how they can benefit from investing in green bonds besides ethical perspectives but considering economic interests. In other words, are green bonds able to bring more profits compared to conventional bonds? With the heating of the green bonds market, various research has been conducted to explore the impact of green premiums. Slimane et al. [2020] gathers all methods that have been applied by most green bond research and fixed effects regression is the most common method. One flaw of this application is time-invariant variables such as dummy variables as a green label cannot fit into the models are therefore will be disregarded. That leaves an empty gap in investigating the pure impact of the green bond label. That is, how much it differs among bonds if it is green labelled. This naturally tosses out one interesting direction to investigate the impact of the green bond label. It is also plausible to hypothesize that the impact of the green label can vary across countries. Factors in terms of politics, resources, the extent of development and so on have a great effect on the evolvement of green bonds. For example, Sweden and Norway have been pioneers and the success factors are largely linked to the support of local governments. The paper lays eyes only on corporate bonds based on the motive that corporates can better represent the demands and changes of the market. Henceforth, the first research question sets the main tone of the paper: *1. Is there any difference in the impact of the green label on the corporates' bonds premium among China, the USA and Europe using the most recent data?*

To use the most recent data, it is inevitable to account for the economic shock caused by the Covid pandemic and lockdown. Some research has studied specifically the performance of sustainable or green bonds in such scenarios. Nofsinger and Varma [2014] find out that mutual funds with ESG factors are more resilient during economic crisis time. Additionally, Silva and Cortez [2016] focus specifically on green bonds and they confirm that green bonds generally do better during an economic crisis. To relate the concurrent

information.

Covid pandemic and the economic shock globally, macro factor as GDP is also taken into account. Therefore, the previous research question is broadened to a bigger picture: *2. What's the impact of GDP on Green bonds? Is the green bond resilient during the Covid pandemic?* To explore the direct relation between GDP and green bonds, an interaction term that combines the green bond label and GDP is imposed and incorporated into the term structure.

It is also introduced earlier that corporate green bonds will be the focus of the paper. It is driven by the motive that corporate bonds are better representatives of market demands and changes. However, corporates differ from each other by their size, industries and so on. To spice the research up, Fama French 3 factor model is renovated and furthermore employed. Fama French 3 factors refer to size premium, value premium and market premium. The assumptions argue that smaller and valued firms tend to give out better premiums. It is therefore natural to suspect the same assumption will resist for green bonds. Henceforth, size premium is complied to investigate the impact of firms' characteristics on green bond premiums. An interaction term that combines green bond label and size premium is incorporated in the term structure as well for exploring the direct link between the two factors. This unfolds the followed-up research question: *3. Will the size of a firm affect the yields of green bonds?*

As aforementioned, many papers show interest in green premium but do not look into the impact of the green bond label directly. This research aims to uncover the direct impact of the green bond label and compare it across countries and regions. Most research incorporates fixed effects where the methods may have default defects in terms of the ignorance of time-invariant variables. In this paper, pooled OLS is employed as a baseline model. However, pooled OLS assumes there is no correlation between time-varying variables and time-invariant variables. This assumption can be hardly fulfilled in practice. For instance, Boutabba and Rannou [2022] argue that green bonds are with higher liquidity risks. To be more explicit, green bonds are less liquid in the secondary market and that causes a comparably larger yield spread. Therefore, the green label is assumed to be correlated with yield spread legitimately. Moreover, Kume and Weir [2013] find out that bond ratings provide informative values and bring a positive impact to yield spreads in a short term. Supposedly, taking bond rating as a latent time-invariant variable and green bond label as an observed time-invariant variable, then it is presumable that yield spread is correlated with both latent time-invariant variables and observed time-invariant variables. This concern is referred to as time-invariant heterogeneity. The econometric model, the Hausman Taylor model, sheds light on tackling this concern. The Hausman Taylor model differentiates time-varying variables that correlated with time-invariant variables and control for the correlations. On account of the fact that the Hausman Taylor model has never been employed in this case before, it is therefore overriding to explore the feasibility of the Hausman Taylor model. The three research questions are evaluated in parallel by both pooled OLS and the Hausman Taylor model. The differences in results are investigated and further analyzed. Moreover, the results are also assessed with the yield curve, where one hypothesizes that the results are able to provide consistent information on the yield curve. Furthermore, backtesting is implemented for validating and comparing the prediction performance of pooled OLS and the Hausman Taylor model. The backtesting is firstly constructed by simulating data from the Garch model. Afterwards, linear predictions of both pooled OLS and the Hausman Taylor model are presented and compared to the Garch forecast. Therefore, the plausibility of the Hausman Taylor model is investigated from both perspectives: the rationality of the results and the prediction performance. By looking into the investigation, one is able to study the influence of time-invariant heterogeneity on the green bond analysis.

The main findings of the research are as follows. Firstly, the yield curves for all countries present an upward trend, which indicates bonds of longer terms give a higher yields payout than the bonds of shorter terms. However, the results of the Hausman Taylor model fail to prevail in any information on the yield curve, whereas pooled OLS provides consistent yield curve information. Moreover, the pooled OLS gives a stronger prediction than the Hausman Taylor and is more aligned with the data simulated from the Garch model. Generating the above information, one concludes that the Hausman Taylor model is outperformed by pooled OLS. This can also be interpreted that time-invariant heterogeneity does not distort the analysis and the pooled OLS offers more plausible findings. Secondly, by interpreting the results from the pooled OLS, the findings are in contrast with Zerbib [2019] in 2019 where the author finds negative greenium. The green bonds that were issued from 2019 to 2021 in the US and Europe are more rewarding than conventional bonds with 0.119 bps and 0.235 bps. Meanwhile, negative greenium is still dominant in China. Thirdly, green bonds in China and Europe are more resilient during the most server economic shock at the beginning of the Covid outburst. The bonds perform generally better during this period and decay when the economy grows back. However, green bonds in the US are always more rewarding than conventional bonds and are not influenced by the economic shift. Moreover, firm size as a factor also affects the green bonds in Europe. It is found that green bonds issued by small firms in Europe give out an extra 0.0401 bps than conventional bonds issued from the same firm.

The remaining parts are organized as follows. section 2 describes recent research and key findings of green bonds. section 3 describes the data construction. section 4 introduces the three hypotheses and two main methods. section 5 documents the empirical results. section 6 reports the robustness of the findings. section 7 concludes the paper.

2 Literature Review

Green bonds have been a trendy topic over the most recent years from both commercial and academic sides. Much research is established to analyze green bonds and the related factors from various angles. Zerbib [2019] presents one of the most inspiring works back to 2019 when he first announce negative greenium. Zerbib [2019] constructs green bond premium by taking the difference between a green bond and a matching counterfactual conventional bond with the same issuer and similar other dynamics. The evidence shows that green bonds premium is 2 basis points generally lower than conventional bonds premium. The difference is significant albeit very small and almost does not differentiate the investors' profits. Moreover, it caters to the preference of investors who are pro-environment. These investors, are oblivious to the negligible difference and would rather hold green bonds instead of conventional bonds with strictly equal risk. The paper was established in 2019 and used the data from July 2013 to December 2017. The conclusion states that the difference between green bonds and conventional bonds is small and even can be omitted. This also suggests that negative basis points from green bond premiums will not affect investors' choices, so to say, scare off investors. The findings provide remarkable insights and attract more eyeballs of the academics. Slimane et al. [2020] contributes to recreating the analysis of greenium Zerbib [2019] while applying two methods: top-down method and bottom-up method. The top-down method refers to comparing a green bond index portfolio to a synthetic conventional green bond index portfolio which is from the same issuer and with a similar maturity. The bottom-up method refers to the matching scheme in that each green bond is matched with a synthetic conventional bond from the same issuer and with the same currency and maturity. The findings are consistent with Zerbib [2019] that both methods give significant and negative greenium of being respectively -4.7 bps and -2.2 bps.

Besides the direct profits obtained from the green bonds, the potential side benefits are also discovered by some papers. Tang and Zhang [2020] analyze green bonds from another perspective by looking into the impact of issuing green bonds on stocks. They find that shareholders gain an advantage when green bonds are issued with the dataset ranging from 2007 to 2017. By issuing green bonds, companies can send the signal to the market that they will make changes to have environmentally friendly investments or projects in the future. This turns out to attract more investors. Stock markets react also positively along with increments in stock price. Moreover, investors will hold the stocks for a comparably longer period instead of realizing within a short time since they believe in a higher valuation of the equity due to the environmentally friendly transformation. However, the positive and promising effects only appear for the first issuance of green bonds but do not happen for the more issuance. Along with the positivity on the stock price brought by the issuance of green bonds, there are other papers that point out that the issuance of green bonds is able to be beneficial for the issuers. Gianfrate and Peri [2019] focus on the green bonds in the European market where they study 121 European green bonds that were issued between 2013 and 2017. Their findings present that by issuing a green-labelled bond, the issuers have lower financial costs in terms of dividends or yields compared to equivalent conventional bonds. This infers that the issuers such as corporates can gather funds for refinancing or investments in a cheaper way.

Research is furthermore extended in the angle to study the characteristic of green bonds under different economic scenarios. One related and inspiring work is established by Nofsinger and Varma [2014], the authors lay eyes on sustainable bonds, where the bonds have profiles in either ethical investing or green investing. In other words, the bonds are highly incorporated with ESG factors(environment, social and governance factors). They find that ESG-driven mutual funds perform worse than conventional mutual funds. However, it outperforms conventional mutual funds during an economic crisis and therefore is more resilient. The findings, therefore, lead to a forward that green bonds may appear to be more resilient than conventional bonds due to their natural characteristics of 'Green'. Similar conclusions have also been reached by Silva and Cortez [2016]. They evaluate bond performance for the US and Europe while taking public and time-varying factors into account. They find that green bonds generally do better during economic crisis periods than in non-crisis periods. More surprisingly, US green funds give back higher yields than other socially responsible funds during the crisis.

Green bonds are nevertheless bonds, that is to say, one cannot ignore the natural characteristics of bonds. Duffie and Singleton [2012] investigate if bond ratings contain pricing-relevant information and they find that rating information does not affect firm value but has an influence on debt value and equity value. Specifically, debt values will increase and equity value will fall when a good rating is given by Moody's and vice versa if a rating is worse than expected. Volatility implied by prices of options from the same issuer will also drop if the issued bond is rated well. Kume and Weir [2013] look further into the relationship between bond ratings and yield spreads. They find out that bond ratings provide investors informative values and the impact of ratings stands out more significantly during economic difficulties. They evidence that there exists a long-term relationship between bond ratings and yield spreads. Moreover, ratings bring a positive impact to yield spreads in a short term. Some studies have also described the link between green bonds and liquidity risk. Boutabba and Rannou [2022] find that green bonds with higher liquidity risks. Therefore, the investors opt for the buy-and-hold strategy. It is also worth noting that the investors are compensated by implied yields from the liquidity risk.

The chapter on green bonds analysis is rather new, that being said, there is much potential room for different methodologies. The methods applied so far in all research do not differ much from each other. Tang and Zhang [2020] use the pooled OLS and employ a dummy variable to indicate a green bond. The defect is pooled OLS assume time-variant variables are uncorrelated with individual time-invariant variables. Whereas Zerbib [2019] take the yield difference between the green bond and the matching conventional bond as a dependent variable and infer there is a green bond premium that affects the yield difference which is time-variant. To take into account the individual time-invariant exogeneity, the authors apply fixed effects to eliminate all individual time-invariant variables. This approach however limits the exploration of the impact of time-invariant variables. Slimane et al. [2020] gather all methods from the recent papers devoted to green bonds, and most of the methods applied to rotate around the above two methods: pooled OLS and fixed effect. As aforementioned, one defect of such methods is the impact of time-invariant variables is neglected. This paper sheds light on exploring the impact of the green bond label directly, yet in the meantime taking individual time-invariant exogeneity into consideration by applying the Hausman-Taylor model. This econometric model is practised with the aim of comparing if the results will differ from OLS or fixed effects. The Hausman Taylor model was firstly imposed by Hausman and Taylor [1981], the primary assumption is that time-variant variables are correlated with time-invariant variables, and further applies the Two-Stage Least-Squares Regression. By such means, time-invariant variables are also being able to be investigated. The Hausman Taylor model has been applied previously with bonds-related topics. Afonso et al. [2007] investigates the factors that determine the sovereign debt credit ratings where one methodology applied is the Hausman Taylor model.

3 Data

This section is divided into a few parts. Firstly, subsection 3.1 describes summary statistics of the bond database of China, the US and Europe. Furthermore, subsection 3.2 dives deeper into mid yield, maturity, rating and average mid yield. Thirdly, subsection 3.3 describes GDP statistics. Finally, subsection 3.4 presents the mid yield difference between big firms and small firms for China, the US and Europe respectively.

3.1 Bond data

The database includes both green bonds and conventional bonds in China, the US and Europe. The bond information is available on Eikon. To account for the currency exchange and different interest rates, only a few countries that use Euro are selected and represented for Europe. Therefore, France, Germany, Spain and Belgium are included in the observation samples for Europe. There are a few selection criteria that have been applied to shape the database. First, only bonds that were issued from 2019-01-01 are considered for the purpose of exploring the most recent green label impact. Secondly, green bonds are selected with the criteria 'Green Bond' as 'Yes', where the GBP are adopted by Eikon. Thirdly, the same bond matching scheme from Zerbib [2019] is employed. More precisely, each green bond is paired with one conventional bond from the same corporate with almost the same issue date (namely, issued within the same year) and the same currency. This naturally leads to a completely balanced bond database where green bonds and conventional bonds share the same population. For the US, there are only 293 green bonds are issued since 2019 which also includes the same issuers. This gives less than 293 corporates for observation. The matching scheme also requires the existence of both green bonds and conventional bonds that are issued within the same year. However, some firms either only issue green bonds that year or the issue date of green bonds and conventional bonds are much apart. Therefore, the number of observed firms is very limited. The same situation also exists in the bonds of China and Europe, where the selection criteria bring down the database size. Henceforth, only around 15 corporates that meet all selection criteria are selected for China, the US and Europe respectively. For each corporate, one green bond and one conventional bond are picked. Totally, there are approximately 30 bonds respectively for China, the US and Europe. Meanwhile, in order to have a fully balanced database that shares the same length of the period for all bonds, the starting points of observation periods for China, the US and Europe are dependent on the latest issued bond date from the respective bond database.

Table 1 presents the size of the database of each country and region as well as the observation period. The data set consists of the following main attributes: bid yield, ask yield, yield change, rating and maturity. It is also worth noticing that the bond data are gathered on a daily basis. There are some missing dates due to the nature of liquidity of the bond market, these missing values are therefore forward refilled by the previous data. For instance, if there are 2 days of data, Day 1 and Day 3, then Day 2 is refilled by Day 1. The full summary statistics for yield spread and mid spread are attached in subsection 8.1.

Country or Region	China	The US	Europe
Number of companies	15	13	12
Number of Green Bonds	15	13	12
Number of Conventional Bonds	15	13	12
Number of total bonds	30	26	24
Observation period	2020-08-07 - 2022-04-21	2021-07-29 - 2022-04-21	2021-02-03 - 2022-04-21

Table 1: General description of the bond database. The starting point observation period depends on the latest issuing date.

3.2 Bonds description

This section inspects the bonds database from several angles: mid price movements, maturity, ratings and average yield. Visualizations are presented to compare and discover if green bonds are particularly distinct from conventional bonds.

3.2.1 Mid yield movements

The first thing to be observed is the mid yield of each bond. Mid yield is the average of ask yield and bid yield. The movement of mid yield gives an idea of the performance and demands of a bond. Figure 3 visualizes mid yield changes over time for selected bonds of China, the US and Europe. Each line stands for the movement of the very one bond during the observation period. It is obvious that most bonds share the same movement patterns, that is, similar ups and downs. The bonds in the US and Europe present an upward-moving pattern whereas the bonds in China are moving slightly downwards. This infers that bonds in the US and Europe are getting more rewarding over time. On the contrary, the bonds in China are paying off less along the time movement. Based on the observations from Europe, there are two outliers that have distinguished higher yields compared to the rest. On the other hand, the spreads of bonds except the outliers do not differ much from each other.

3.2.2 Maturity

As previously mentioned, green bonds are mostly issued for the purpose of refinancing for sustainable projects. On that account, it is legitimate to suspect that green bonds with longer maturity. To identify if there indeed exists such a unique characteristic in the database, an overview summary of maturities is visualized for easier comparison. Figure 4 shows that from selected Chinese bonds, green bonds are absent from one year of maturity. Meanwhile, green bonds are slightly more active in maturity being 3 years and 5 years, whereas conventional bonds have more shares in longer maturity being 7 years. The American bonds sample display that green bonds are more varied compared to conventional bonds, where the former appear in all maturities. Lastly, green bonds show almost equally active in all maturities as conventional bonds for the selected European samples. To conclude, there is no particularly prominent difference in maturities distribution between green bonds and conventional bonds for all observed countries.

3.2.3 Ratings

It is also intriguing to see if ratings will differ within green-labelled and non-green labelled bonds. The most commonly adopted rating standards are Moody's or S&Ps in both the US and Europe. However, these two rating methodologies do not cover elaborately the bonds in China. There are 11 credit domestic agencies totally in China, and the rating for a bond differs in various agencies. In other words, finding consistency in ratings for all bonds that share the same standards or from the same agency is rather tricky. It is also tedious to convert the ratings that are in line with the standards of S&P or Moody's. Therefore, incomplete rating information is disregarded in this paper and treated as an unobserved time-invariant variate.

Nevertheless, the bond rating information for the US and Europe is obtained from S&P and furthermore analyzed to explore the direct link between rating and green bond label from a superficial level. There is a total of 17 levels in ratings, ranging from AAA to D, based on the descending repayment probability. The full chart of S&P levels and according to repayment description is attached in section 8.

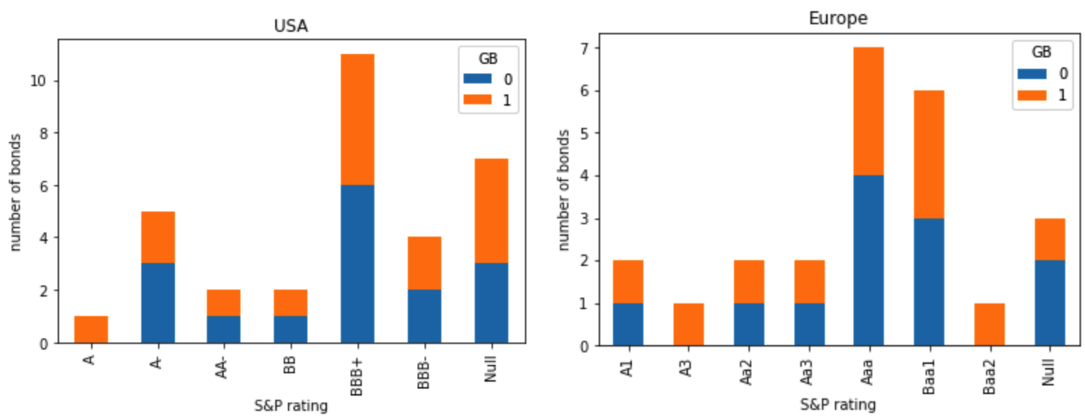
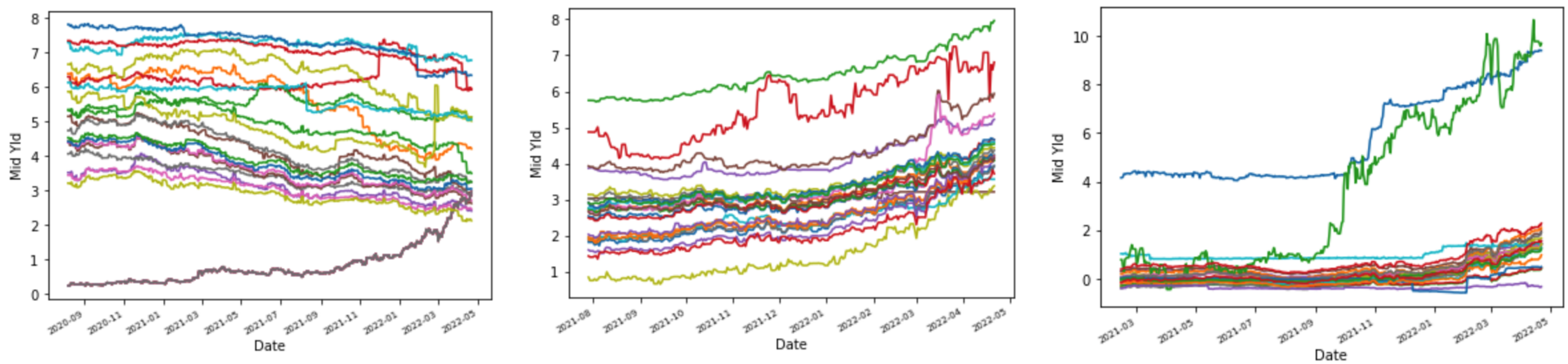


Figure 2: Overview of S&P ratings for selected USA bonds and Europe bonds.

Figure 2 shows that ratings do not differ much in green labels. Both green bonds and conventional bonds are almost equally distributed in all rating levels for the US and Europe. Except that in the US, there are A3 and Baa2 rated green bonds whereas conventional bonds are absent from these two ratings. Slimane et al. [2020] has a similar analysis with maturities and ratings that incorporate a larger database that consists of 532 green bonds and a total of around 1000 bonds. The author discovers that green bonds have higher shares in the bucket of maturity of 5 years and 10 years compared to conventional bonds. As for ratings, the distribution is generally well balanced except green bonds populated more in Aa and Baa ratings. These findings are consistent with this paper and provide a broader picture where the bonds database has a longer observation period starting from 2016.

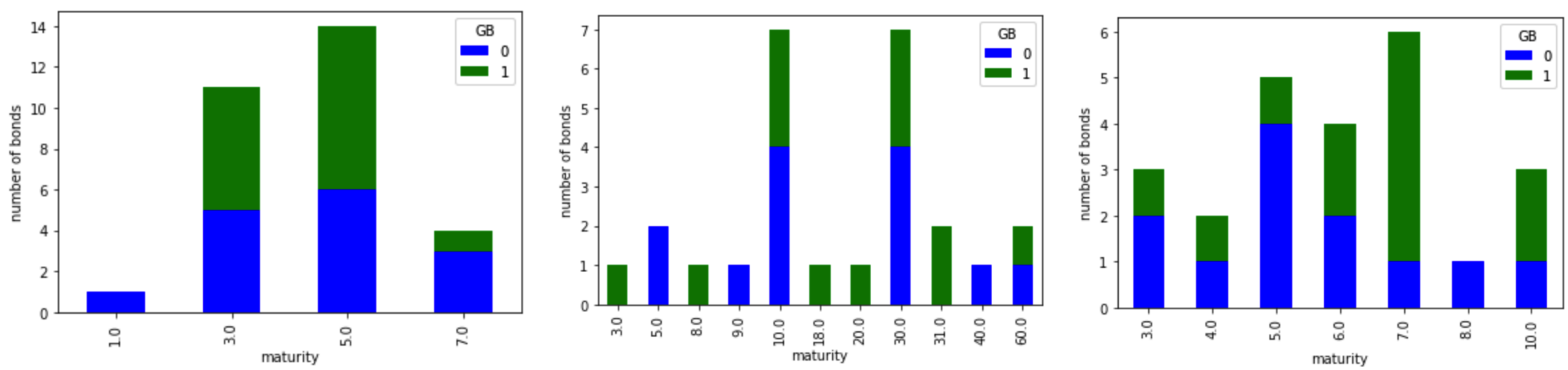


(a) Mid yield of bonds in China with CNY

(b) Mid yield of bonds in the US with USD

(c) Mid yield of bonds in Europe with EUR

Figure 3: Mid yield over time for China, the US and Europe selected bonds.



(a) China

(b) USA

(c) Europe

Figure 4: Overview of maturity for selected bonds from China, the US and Europe. Label 1 stands for the green bond and 0 stands for the conventional bond. Note that the maturities appear distinctively in China, the US and Europe. In China, only 1 year, 2 years, 5 years and 7 years are active. The maturities in the US are the most extensive and then followed by Europe.

3.2.4 Average yield

Lastly, the spotlight is put on the average yield. Bonds are separated into two groups: green bonds and conventional bonds. By comparing the average yields of green bonds and conventional bonds, one can detect the general performance of bonds with different labels. Figure 5 presents the movements of average yields for China, the US and Europe within the observed period. Interestingly, green bonds in China are more rewarding than conventional bonds before 2021-07. However, the outperformance decays afterwards and is gradually taken over by conventional bonds. It is also worth noticing the overall downward moving trend is significantly opposite to the US and Europe. For the USA, the first thing to be noted is that green bonds give out a higher payout at almost all times than conventional bonds. Additionally, the difference is not significantly big. Moreover, the average yield for both green bond and non-green bonds have almost the same moving patterns and are increasing over time. Moving to Europe, it is outstanding again that the average yields of green bonds are generally higher than that of conventional bonds during the whole observed period. The difference is much bigger at the early stage and it shrinks over time. Based on Silva and Cortez [2016], green bonds generally have a better performance during crisis times. Therefore, one plausible inference can be the special prevalence of green bonds during economic crises fades away gradually due to the revival of economics. These superficial eyeballing observations are not surprising and can be referred to Nofsinger and Varma [2014] and Silva and Cortez [2016] where they find that green mutual funds are more resilient during economic crises compared to conventional mutual funds.

3.3 Macro data: GDP

GDP(Gross Domestic Product) stands for the total market value of all the final goods and services produced in the year by the country. It is one of the most important monetary measures for economic growth. The information on GDP is provided by *The World Bank*[2] and *China Data Online*[3]. To present economic growth, the percentage of the growth of the current year compared to the previous year is provided in Figure 6. The overview statistics are presented in subsection 8.2 in Appendix.

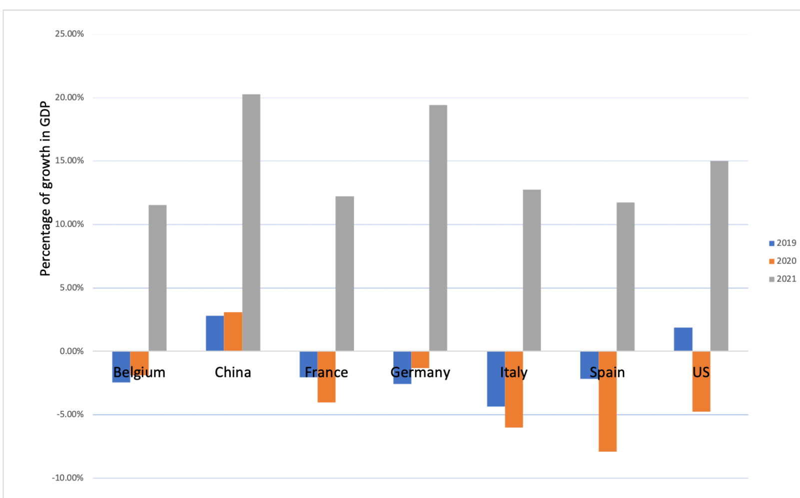
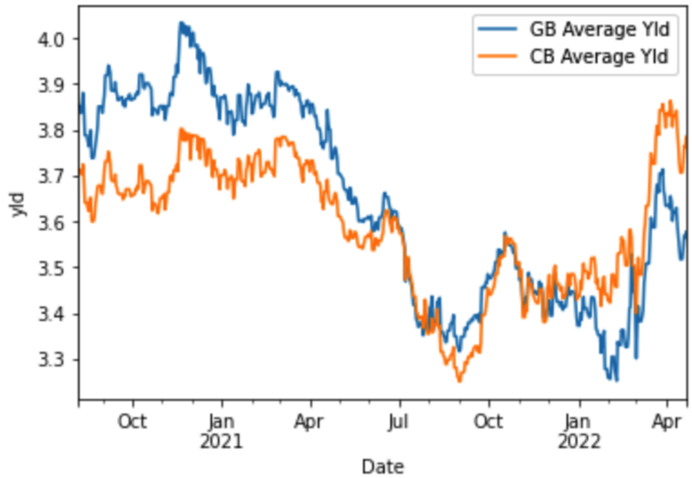
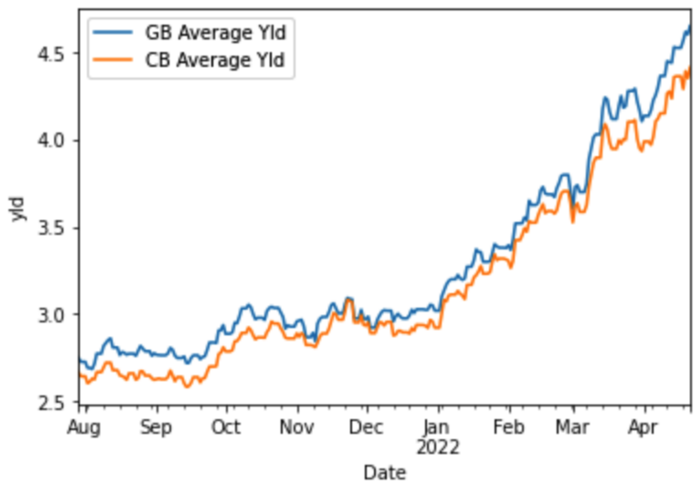


Figure 6: GDP growth from 2019 to 2022 for observed countries.

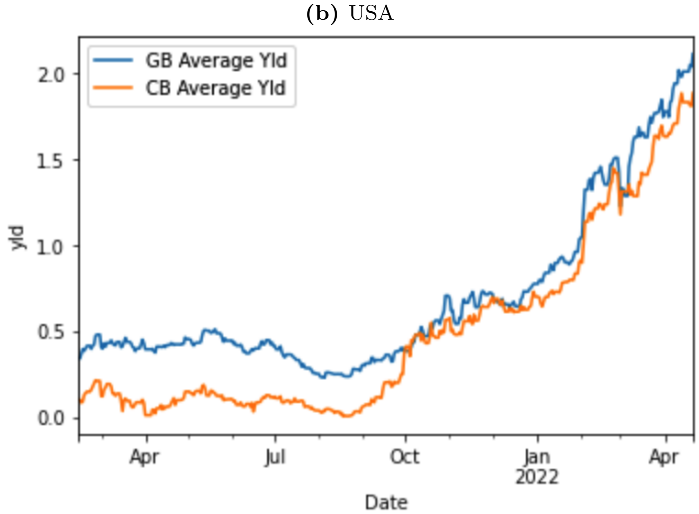
[2]https://data.worldbank.org/indicator/SL.UEM.TOTL.ZS
[3]https://www.china-data-online.com/member/macroy/macroyadv.asp



(a) China



(b) USA



(c) Europe

Figure 5: Average yield for green bonds and conventional bonds for China, the US and Europe.

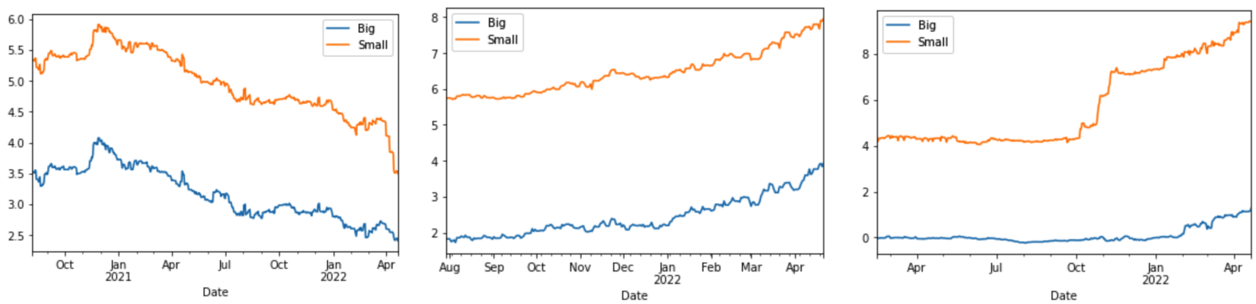
Figure 6 illustrates that GDP growth for all observed countries from 2019 to 2021. It is not surprising that most countries suffer from negative growth due to the outburst of Coronavirus in 2019 and 2020 except for China and the US. In the year 2021, the growth is positive again with great improvement. This also can be reflected by the releasing of lockdown and reviving of economics.

3.4 Size premium

Fama French 3 factor model is one famous asset pricing model that explains the stock returns with 3 factors. Fama French 3 factors consist of value premium(HML), size premium(SMB) and excess premium of a market portfolio that excludes risk-free rate. The value premium is referred to as high minus low, which is also referred to as the spread in returns between values stocks and growth stocks. Size premium stands for small minus big, which is shorthand for the return spread between small firms and big firms. Lastly, the excess premium of a market portfolio is the return from the benchmark market portfolio index. The model is widely used in analyzing the equity markets, such as stocks. However, the model is applied and extended paralleling for credit markets. Bektić et al. [2019] examines the pricing impact of the Fama French 4 factors, size, value, profitability and investment in the US and European corporate bond market. Similar studies have also been conducted by Houweling and Van Zundert [2017], where the authors examine the corporate bond market with 4 risk factors: size, low-risk, value and momentum. One would naturally suspect if these risk factors, specifically, size and value could play any effects on the bond yields that are interested in this paper. One essential step is to identify these two factors. The most common measure for identifying the size of firms is to use of market capitalization. It is calculated as multiplying the current share price by the number of outstanding shares. However, market capitalization varies monthly or even daily due to the volatile nature of stock. Shareholder equity on the other hand is a statement of a company's assets minus its liabilities and it does not fluctuate based on the stock price. Therefore shareholder equity is preferred as the measure for the size factor in this paper. The information of equity data is extracted from the balance sheet of 2020-12-30 on **Bloomberg**. Size premium is henceforth calculated as the yield difference between the firms with the highest equity and the lowest equity. Furthermore, the most common measure for value stocks is a high book-to-market ratio. Nonetheless, some firms in the observations lack stock information or there is no up-to-date balance sheet, one is restricted to take market capitalization into account and therefore cannot further incorporate with value premium. For that reason, only size premium is comfortable to be examined in this paper and it is constructed as:

$$\text{SMB} = \text{Mid Yld}_{Small} - \text{Mid Yld}_{Big} \tag{1}$$

Fama French 3 factor model assumes that smaller firms reward better returns than bigger firms. Figure 7 confirms the assumption and displays that small firms generally give out higher yields than big firms.



(a) China **(b)** USA **(c)** Europe

Figure 7: Mid yield for big and small firms for China, the US and Europe separately.

4 Hypotheses and Methodology

As introduced in section 1, the paper intends to answer the main three research questions. That is, the impact of the green bond label, the GDP factor and size premium respectively and the relation of the two latter factors with the green bond label. Corresponding hypotheses for individual research questions are first proposed. To examine the hypotheses, methods such as pooled OLS and the Hausman Taylor model are implemented and renovated. The section is constructed as follows: first, subsection 4.1 proposes the hypotheses. Then, subsection 4.2 introduce main methods such as pooled OLS and the Hausman Taylor model in detail. subsection 4.3 puts forwards the methods to examine the robustness of the findings. subsubsection 4.3.1 introduces the variance inflation factor for the aim of a multicollinearity check. subsubsection 4.3.2 proposes different variable setups for the Hausman Taylor model to inspect the difference in results. Lastly, subsubsection 4.3.3 presents the Garch model which is employed to simulate data for backtesting.

4.1 Hypotheses

4.1.1 Hypothesis 1: Baseline Model

Recall the first research question, 'Is there any difference in the impact of the green label on the corporates bonds premium among China, the USA and Europe using the most recent data?' To inspect the impact of the green bond label, a regression analysis is constructed with employing the green bond label as a time-invariant explanatory variable. Along with the green bond label, another time-invariant variable included is maturity. As for time-varying dynamics, bid yield change and yield spread are employed based on the fact that they vary daily. Moreover, mid yield is treated as a dependent variable. The hypothesis states that mid yields are correlated and can be explained mainly by these four named explanatory variables. The linear regression is constructed as:

$$\text{Mid yield}_{it} = \beta_0 + \beta_1 \text{Bid yield change}_{it} + \beta_2 \text{Yield spread}_{it} + \delta_1 \text{Green bond}_i + \delta_2 \text{Maturity}_i + \epsilon_{it}, \tag{2}$$

where β_i and δ_i are coefficients for time-varying variables and time-invariant variables respectively, β_0 is a constant and ϵ_{it} is the individual error term. However, besides the observed time-invariant variables such as maturity and green bond, there is also latent time-invariant variables such as ratings, industries and so on are not included due to the difficulty of data gathering. One alleged assumption is some of the variants are potentially correlated with unobserved time-invariant variables. Duffie and Singleton [2012] and Kume and

Weir [2013] indicate that ratings provide pricing-related information and moreover have a positive impact on yield spread. Henceforth, it is likely that yield spread is correlated with ratings, where the former is referred to as time-varying and the latter is referred to as an unobserved time-invariant variable. Meanwhile, it is also plausible to hypothesise that green bonds tend to have a higher rating based on the fact that most projects are facilitated for sustainability and are supported by governments. Therefore, time-varying variables as yield spread and time-invariant variables such as the green bond label are suspected to be correlated with unobserved individual effects.

4.1.2 Hypothesis 2: Incorporation of Macro variable

To answer the second research question, that is, '2. What's the impact of GDP on Green bonds? Is the green bond resilient during the Covid pandemic?', the baseline model introduced earlier is therefore extended with the GDP factor. To translate literally, the hypothesis believes that mid yield can be explained by green bond label, maturity, bid yield change, yield spread and the newly added GDP factor. The regression is unfolded as:

$$\text{Mid yield}_{it} = \beta_0 + \beta_1 \text{Bid yield change}_{it} + \beta_2 \text{Yield spread}_{it} + \beta_3 \text{GDP}_{it} + \delta_1 \text{Green bond}_i + \delta_2 \text{Maturity}_i + \epsilon_{it}. \tag{3}$$

With the new variable GDP intervened, it is again vital to identify if it is correlated with the unobserved time-invariant variable. However, in this case, GDP is observed from a country level and is therefore not assumed to be correlated with individual effects. This subsequently leaves the same presumption for unobserved heterogeneity, that is, green bond label and yield spread are potentially correlated with latent variables only.

Furthermore, Nofsinger and Varma [2014] and Silva and Cortez [2016] discuss the resilience of green bonds during economic crisis time. Therefore, the direct relation between green bond label and GDP is also under investigation. The implementation can be fulfilled by creating an interaction term which is the product of the green bond label and GDP factor. Henceforth, the interaction term is included in the aforementioned GDP extended model. The equation is thus renovated as:

$$\begin{aligned} \text{Mid yield}_{it} = \beta_0 &+ \beta_1 \text{Bid yield change}_{it} + \beta_2 \text{Yield spread}_{it} + \beta_3 \text{GDP}_{it} + \beta_4 \text{GDP}_{it} * \text{Green bond}_i \\ &+ \delta_1 \text{Green bond}_i + \delta_2 \text{Maturity}_i + \epsilon_{it}. \end{aligned} \tag{4}$$

4.1.3 Hypothesis 3: Incorporation of the size factor

The third research question involves the firm-specific characteristics, namely, the firm size. For that reason, the size factor is extended based on the baseline model. The hypothesis presumes that mid yield can be described by bid yield change, yield spread, green bond label, maturity and size factor. The equation thus follows:

$$\text{Mid yield}_{it} = \beta_0 + \beta_1 \text{Bid yield change}_{it} + \beta_2 \text{Yield spread}_{it} + \beta_3 \text{Size premium}_{it} + \delta_1 \text{Green bond}_i + \delta_2 \text{Maturity}_i + \epsilon_{it}. \tag{5}$$

The same assumption for unobserved-heterogeneity from the earlier section is also applied here where green bond and yield spread are correlated with latent individual effects. What's more, there is hardly any inference that the size factor is correlated with any latent unobserved variables. Therefore, size premium is taken as exogenous.

Again, it arises the curiosity about the direct correlation between the green bond label and the size factor. The same procedure is hence implemented, an interaction term combines the green bond label and the size factor is incorporated in the aforementioned size factor expanded model. The model, therefore, is developed as:

$$\begin{aligned}\text{Mid yield}_{it} &= \beta_0 + \beta_1\text{Bid yield change}_{it} + \beta_2\text{Yield spread}_{it} + \beta_3\text{Size premium}_{it} \\ &+ \beta_4\text{Size premium}_{it} * \text{Green bond}_i + \delta_1\text{Green bond}_i + \delta_2\text{Maturity}_i + \epsilon_{it}.\end{aligned} \tag{6}$$

4.2 Methodology

4.2.1 Method 1: Pooled OLS

To proceed and further examine the aforementioned hypotheses, two methods such as pooled OLS and the Hausman Taylor model are employed. Slimane et al. [2020] summarize all the methods that have been adopted for green bonds research and the most frequent methods are to use pooled OLS and fixed effects. However, the fixed effect is limited in the paper due to the reason that time-invariant variables are eliminated. Moreover, there are embedded defects in pooled OLS and therefore the Hausman Taylor model is proposed for further practice. In this subsection, pooled OLS and the Hausman Taylor model are established in detail and renovated individually for different hypotheses.

Pooled OLS is given as follows:

$$Y_{i,t} = X_{it}\beta + Z_i\delta + \epsilon_{it} \quad \text{for t=1,...,T and i=1,...,N,} \tag{7}$$

where X_{it} is time-varying variables and Z_i is observed time-invariant variables. To link with the aforementioned hypotheses and pratice with pooled OLS, Table 2 presents the full overview of the allocation of the variables.

	X_{it}	Z_{it}
Hypothesis 1	$\text{Bid yield change}_{it}, \text{Yield spread}_{it}$	$\text{Green bond}_i, \text{Maturity}_i$
Hypothesis 2	$\text{Bid yield change}_{it}, \text{Yield spread}_{it}, \text{GDP}_{it}$	$\text{Green bond}_i, \text{Maturity}_i$
Hypothesis 2*	$\text{Bid yield change}_{it}, \text{Yield spread}_{it}, \text{GDP}_{it}, \text{GDP}_{it} * \text{Green bond}_i$	$\text{Green bond}_i, \text{Maturity}_i$
Hypothesis 3	$\text{Bid yield change}_{it}, \text{Yield spread}_{it}, \text{Size premium}_{it}$	$\text{Green bond}_i, \text{Maturity}_i$
Hypothesis 3*	$\text{Bid yield change}_{it}, \text{Yield spread}_{it}, \text{Size premium}_{it}, \text{Size premium}_{it} * \text{Green bond}_i$	$\text{Green bond}_i, \text{Maturity}_i$

Table 2: Sets for time-invariant and time-varying variables for pooled OLS for hypotheses. * refers to the model extended with an interaction term.

Moreover, the assumption $cov(X_{it}, Z_i) = 0$ holds. This assumption ignores the potential correlation between individual time-variant variables and latent time-invariant variables, which is also called time-invariant heterogeneity. It also ignores that Z_i has potential serial correlation over time. To illustrate unobserved time-heterogeneity, suppose there is an unobserved time-invariant variable u_i. The linear model,Equation 7 therefore, is adapted to:

$$Y_{i,t} = X_{it}\beta + Z_i\delta + u_i + \epsilon_{it} \quad \text{for t=1,...,T and i=1,...,N.} \tag{8}$$

This further provides:

$$Y_i= \begin{pmatrix} Y_{i,1} \\ Y_{i,2} \\ \vdots \\ Y_{i,T} \end{pmatrix} = \begin{pmatrix} X_{i1}\beta + Z_i\delta + u_i + \epsilon_{i1} \\ X_{i2}\beta + Z_i\delta + u_i + \epsilon_{i2} \\ \vdots \\ X_{iT}\beta + Z_i\delta + u_i + \epsilon_{iT} \end{pmatrix} = \begin{pmatrix} X_{i1} & Z_i \\ X_{i2} & Z_i \\ \vdots & \vdots \\ X_{iT} & Z_i \end{pmatrix} \begin{pmatrix} \beta \\ \delta \end{pmatrix} + \begin{pmatrix} u_i \\ u_i \\ \vdots \\ u_i \end{pmatrix} + \begin{pmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \vdots \\ \epsilon_{iT} \end{pmatrix}.$$

Since both X_{it} and Z_i are observed variables. For the sake of simplicity, X_{it} and Z_i are combined as M_{it} with the vector η. Gather all individuals and present it in a matrix form:

$$Y_i = M\eta + U + \epsilon, \tag{9}$$

with M= $\begin{pmatrix} X_{i1} & Z_i \\ X_{i2} & Z_i \\ \vdots & \vdots \\ X_{iT} & Z_i \end{pmatrix}$, $\eta = \begin{pmatrix} \beta \\ \delta \end{pmatrix}$, U= $\begin{pmatrix} u_i \\ u_i \\ \vdots \\ u_i \end{pmatrix}$, $\epsilon = \begin{pmatrix} \epsilon_{i1} \\ \epsilon_{i2} \\ \vdots \\ \epsilon_{iT} \end{pmatrix}$ and the expected value of U is 0.

The estimator of η solved by Ordinary Least Squares is:

$$\hat{\eta} = (M'M)^{-1}M'Y, \tag{10}$$

where U contains unobserved variables and is omitted by the OLS estimator. Substitute Y_i back to get:

$$\begin{aligned} \hat{\eta} &= (M'M)^{-1}M'(M\eta + U + \epsilon) && (11) \\ &= (M'M)^{-1}M'M\eta + (M'M)^{-1}M'U + (M'M)^{-1}M'\epsilon && (12) \\ &= \eta + (M'M)^{-1}M'U + (M'M)^{-1}M'\epsilon. && (13) \end{aligned}$$

After taking the expectation conditional on M:

$$E(\hat{\eta}|M) = \eta + (M'M)^{-1}E(M'U|M), \tag{14}$$

where $E(\epsilon|M) = 0$ is assumed. The second term is not eliminated due to the observed variable is correlated with unobserved variables, that is, M is correlated with U. This term is therefore referred to as bias that causes the inefficiency of the OLS estimator.

To remedy the bias brought by time-invariant heterogeneity, the most common solution is to use fixed-effect regression by taking the first differences using the average from both sides. The illustration shows:

$$\begin{aligned} Y_{it} - \bar{Y}_i &= \beta_i(X_{it} - \bar{X}_i) + \delta_i(Z_i - \bar{Z}_i) + (u_i - \bar{u}) + (\epsilon_{it} - \bar{\epsilon}_i) && (15) \\ \tilde{Y}_{it} &= \beta_i\tilde{X}_{it} + \tilde{\epsilon}_{it}. && (16) \end{aligned}$$

However, it also eliminates the time-invariant variables Z_i which are part of the main focus. The aforementioned paper such as Zerbib [2019] employs fixed effects regression where all constant variables are eliminated.

4.2.2 Method 2: The Hausman Taylor Model

The preceding section explains that pooled OLS can be distorted by time-invariant heterogeneity. Hausman and Taylor [1981] shed light on exploring the impact of individual time-invariant variables while considering unobserved time-invariant variables. The procedures are summarized by Ao [2009] and the essence is given as follows:

$$Y_{it} = X_{it}\beta + Z_i\delta + u_i + \epsilon_{it}, \tag{17}$$

where X_{it} are time-varying variables, Z_i are time-invariant variables, u_i is the latent time-invariant variables and ϵ_{it} is idiosyncratic error term with the mean of 0. X_{it} and Z_{it} are then separated into 2 sets, $X_{it} = [X_{1it}, X_{2it}]$ and $Z_{it} = [Z_{1it}, Z_{2it}]$. The assumptions followed with:

$$cov(X_{1it}, u_i) = 0 \tag{18}$$

$$cov(Z_{1it}, u_i) = 0 \tag{19}$$

$$cov(X_{2it}, u_i) \neq 0 \tag{20}$$

$$cov(Z_{2it}, u_i) \neq 0. \tag{21}$$

Replace X_{it} and Z_{it} with the sets to get:

$$Y_{it} = X_{1it}\beta_1 + X_{2it}\beta_2 + Z_{1i}\delta_1 + Z_{2i}\delta_2 + u_i + \epsilon_{it}. \tag{22}$$

For the first step, within regression which is also referred as Equation 15 is examined to get the coefficients of the time-varying variables:

$$\tilde{Y}_{it} = \tilde{X}_{1it}\hat{\beta}_{1w} + \tilde{X}_{2it}\hat{\beta}_{2w} + \tilde{\epsilon}_{it}, \tag{23}$$

where $\hat{\beta}_{1w}$ and $\hat{\beta}_{2w}$ are the coefficients that are asked for. The next step is to obtain "within" residual:

$$\tilde{d}_{it} = \tilde{y}_{it} - \tilde{X}_{1it}\hat{\beta}_{1w} - \tilde{X}_{2it}\hat{\beta}_{2w}, \quad \text{and this gives} \quad \hat{\sigma}^2_\epsilon = \frac{\sum_t \sum_i \tilde{d}_{it}}{N-n}. \tag{24}$$

Furthermore, regress $\tilde{d}_{it}$ on Z_1 and Z_2 where X_1 and Z_1 are taken as instrumental variables respectively. The coefficients of Z_1 and Z_2 are obtained as $\tilde{\delta}_{1IV}$ and $\tilde{\delta}_{2IV}$, where $\tilde{\delta}_{IV}$ is the in the form of: $\tilde{\delta}_{IV} = (Z'P_AZ)^{-1}Z'P_A\tilde{d}$, with $P_A = A(A'A)^{-1}A'$ and $A = [X_1, Z_1]$. With the information of $\tilde{\delta}_{IV}$ and $\hat{\sigma}_\epsilon$, one is able to obtain $\hat{\sigma}^2_u$. Finally, combine both $\hat{\sigma}^2_\epsilon$ and $\hat{\sigma}^2_u$ to get: $\hat{\theta}_i = 1-(\frac{\hat{\sigma}^2_\epsilon}{\hat{\sigma}^2_\epsilon + T\hat{\sigma}^2_u})^{\frac{1}{2}}$. The Gauss-Markov estimator, which is the minimum variance matrix-weighted average of the within and between groups estimators is then acquired:

$$w^* = [X_{1it}, X_{2it}, Z_{i1}, Z_{2i}] - \hat{\theta}_i[X_{1it}, X_{2it}, Z_{i1}, Z_{2i}]. \tag{25}$$

To generate also Y^* and V_{it}^*:

$$Y^* = Y_{it} - \hat{\theta}Y_{it}, V_{it}' = [(X_{1it} - \bar{X}_{1i})',\ (X_{2it} - \bar{X}_{2i})', Z_{1i}', \bar{X}_{1i}']. \tag{26}$$

Then examine a 2SLS regression of Y^* and W^* with V_{it} being the instruments to get the desired coefficients of ($\hat{\beta}'$, $\hat{\alpha}'$).

To apply the Hausman Taylor model for the proposed hypotheses, the essential step is to differentiate x_1, x_2, z_1 and z_2. Table 3 displays the allocation for the variables.

	X_1	X_2	Z_1	Z_2
Hypothesis 1	Bid yield Change_{it}	Yield spread_{it}	Green bond_i	Maturity_i
Hypothesis 2	Bid yield $\text{Change}_{it}, \text{GDP}_i$	Yield spread_{it}	Green bond_i	Maturity_i
Hypothesis 3	Bid yield Change_{it}, Size premium_{it}	Yield spread_{it}	Green bond_i	Maturity_i

Table 3: Sets for X_1, X_2, Z_1 and Z_2 for the Hausman Taylor model for hypotheses.

4.3 Robustness check

4.3.1 Multicollinearity check: Variance inflation factor

Multiple linear regressions are throughout the whole analysis, which is interpreted as multiple explanatory variables incorporated in the linear regression. One assumption of multiple linear regression is that the explanatory variables are independent of each other. The counter case where explanatory variables are correlated to each other is known as multicollinearity. Multicollinearity brings difficulties to examine the individual effect of each explanatory variable considering the fact that the effects are overlapped. Therefore, it is essential to check first that there is no multicollinearity. The Variance inflation factor(VIF) is thus employed for checking multicollinearity. Consider the linear model:

$$Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k + \epsilon.$$

The estimated variance of estimate β_j can be expressed as:

$$\hat{\text{var}}(\hat{\beta}_j) = \frac{s^2}{(n-1)\hat{\text{var}}(X_j)} \cdot \frac{1}{1-R_j^2},$$

where s is the root mean squared error, n is the sample size, R^2 is the percentage of variance from the dependent variable that is explained by the explanatory variable with the range varying from 0 to 1. The focus is the term $\frac{1}{(1-R_j^2)}$ which is also referred to as VIF. The higher value of R^2, the higher value of VIF, which also means a higher correlation within explanatory variables. One can firstly run the linear regression with only explanatory variables involved:

$$X_i = \alpha_0 + \alpha \mathbf{X}_{-i} + e,$$

where $\mathbf{X}_{-i}$ is the rest explanatory variables excluding X_i, α_0 is a constant and e is the error term. Then VIF can be calculated as: $\text{VIF}_i = \frac{1}{1-R_i^2}$. The degree of multicollinearity can be counted to be high when $\text{VIF}(\hat{\beta}_i)$

> 10. Meanwhile, if VIF_i is 1, it hints that X_i is diagonal with the rest of the explanatory variables.

4.3.2 Twisting X and Z for the Hausman Taylor Model

Earlier it is introduced that in the setup of the Hausman Taylor model, X and Z are both split into two sets, one set is correlated with the unobserved time-invariant variables and the other one is not. Table 3 nominates the allocation of variables for different hypotheses. However, the allocations are implemented under assumptions, that being said, they are not supported by ample evidence. Therefore, one proposes to twist the variables and inspect if there are any major changes in the results. It is sceptical if the green label truly relates to the internal features of bonds. Henceforth, the first twist experiments on the green bond label, where the green bond label is treated as exogenous instead.

	X_1	X_2	Z_1	Z_2
Hypothesis 1	Bid yield Change$_{it}$	Yield spread$_{it}$	Maturity$_i$	Green bond$_i$
Hypothesis 2	Bid yield Change$_{it}$, GDP$_i$	Yield spread$_{it}$	Maturity$_i$	Green bond$_i$
Hypothesis 3	Bid yield Change$_{it}$, Size premium$_{it}$	Yield spread$_{it}$	Maturity$_i$	Green bond$_i$

Table 4: Sets for X_1, X_2, Z_1 and Z_2 for the Hausman Taylor model for hypotheses, with Green bond$_i$ is treated as exogenous.

The second twist happens on GDP$_i$ and Size premium$_i$ where the two variables enter as exogenous previously. The same suspicion arises: if the empirical results will change when these two factors are treated as endogenous, or in other words if they are assumed to be correlated with latent time-invariant variables. Thus one employs these two variables as endogenous for the twisting test.

	X_1	X_2	Z_1	Z_2
Hypothesis 2	Bid yield Change$_{it}$	GDP$_i$, Yield spread$_{it}$	Green bond$_i$	Maturity$_i$
Hypothesis 3	Bid yield Change$_{it}$	Size premium$_{it}$, Yield spread$_{it}$	Green bond$_i$	Maturity$_i$

Table 5: Sets for X_1, X_2, Z_1 and Z_2 for the Hausman Taylor model for hypotheses with GDP$_i$ and Size premium$_{it}$ treated as endogenous respectively for Hypothesis 2 and Hypothesis 3.

4.3.3 Backtesting: The Garch Model

To compare the prediction power of two models, pooled OLS and the Hausman Taylor model. It is an essential step for backtesting. Backtesting is the method that applies a series of unconsumed data over the methodologies to get predictions and compare the predictions with the realized outcome. Therefore, a series of simulated data that captures all characteristics of the original data is demanded for backtesting. The Garch model is therefore employed. The Garch model is referred to as Generalized AutoRegressive Conditional Heteroskedasticity. The core idea is that volatility today depends on previous error terms and previous volatility. In other words, one can predict the volatility today if the previous error terms and previous volatility are known. The model gives as follows:

$$y_t = x_t^{'} b + \epsilon_t, \tag{27}$$

$$\epsilon_t | \psi_{t-1} \sim \mathcal{N}(0, \sigma_t^2), \tag{28}$$

$$\sigma_t^2 = w + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2, \tag{29}$$

where in Equation 29, ϵ_{t-1} is the error term at t-1 and σ_{t-1} is volatility at time t-1. p and q stand for the number of previous terms for error terms and volatility. In this research, p and q are both taken as 2. Equation 29 generates the prediction of σ_t at present, which further can be employed to predict ϵ_t in Equation 28. The demanded estimation of y is thus followed by Equation 27 where ϵ_t is obtained from earlier.

5 Results and Analysis

This section reports the empirical results for both pooled OLS and the Hausman Taylor model for all hypotheses. To start with, subsection 5.1 describes the results regarding to the baseline model, which is also referred as the first hypothesis. Next, subsection 5.2 presents the results for the second hypothesis. Furthermore, subsection 5.3 states the results for the third hypothesis. Lastly, subsection 5.4 combines all data from all countries and regions together and experiment the combined data on the first hypothesis. This sub section contributes to investigate the impact of the green bond label in a more universal and broader perspective.

5.1 Hypothesis 1: Baseline model

This section documents the empirical results of the first hypothesis subsubsection 4.1.1, which is also referred to as the baseline model. Two methods as pooled OLS and the Hausman Taylor model are implemented. The latter accounts for the potential time-invariant heterogeneity.

	China		US		Europe	
VARIABLES	Pooled OLS	HT	Pooled OLS	HT	Pooled OLS	HT
Yld Spread	42.16***	16.40***	8.126***	39.70***	21.21***	45.12***
(TVendogenous)	(0.208)	(0.391)	(0.236)	(1.285)	(0.237)	(0.380)
Bid Yld Change	1.178***	1.065***	3.708***	3.287***	-0.000744	0.000681
(TVexogenous)	(0.283)	(0.126)	(0.345)	(0.251)	(0.00181)	(0.000863)
Green Bond	-0.137***	19.35	0.119***	-3.448	0.235***	-2.438223
(TIendogenous)	(0.0170)	(35.93)	(0.0170)	(4.941)	(0.0236)	(3.6942)
Maturity	0.282***	0.749	0.0580***	0.0729**	0.208***	0.5272*
(TIexogenous)	(0.00591)	(1.329)	(0.000557)	(0.0292)	(0.00600)	(0.307)
Constant	-0.526***	-10.51	1.335***	0.686	-2.185***	-4.3574**
	(0.0295)	(21.15)	(0.0250)	(2.488)	(0.0441)	(1.4090)
Observations	18,690	18,690	6,942	6,942	10,392	10,392
Number of Bonds	30	30	26	26	24	24
R-squared	0.731		0.620		0.443	

***, ** Statistically significant at the 1% level and 5% level, respectively.

Table 6: Empirical results of Pooled OLS and Hausman-Taylor methods for baseline model to investigate green bonds labels. (OLS: $\text{Mid Yld}_{it} = \beta_0 + \beta_1 \text{BidyChg}_{it} + \beta_2 \text{Yld spread}_{it} + \delta_1 \text{GB}_i + \delta_2 \text{Maturity}_i + \epsilon_{it}$, HT: $\text{Mid Yld}_{it} = \beta_0 + \beta_1 \text{BidyChg}_{it} + \beta_2 \text{Yld spread}_{it} + \delta_2 \text{GB}_i + \delta_1 \text{Maturity}_i + u_i + \epsilon_{it}$, with $[\mathbf{x}'_{1it}] = [\text{BidyChg}_{it}]$, $[\mathbf{x}'_{2it}]=[\text{Yld spread}_{it}]$ $[\mathbf{z}'_{1it}]=[\text{Maturity}_i]$ and $[\mathbf{z}'_{2it}]=[\text{GB}_i]$.) Table 6 reports the regression results for China, the US and Europe respectively. The starting date for each bond is dependent on the lasted issued bond, this leads to the same length of the bond database. The database period is 2020-08-07 - 2022-04-22, 2021-07-29 - 2022-04-22 and 2020-12-15 - 2022-04-22 respectively.

Table 6 reports that in pooled OLS models, there exists negative greenium for China with the value of -0.137. The results are consistent with Zerbib [2019] where he also found negative greenium. However, it is not the case for the US and Europe where the green bond label has a positive and significant impact of 0.119 and 0.235 respectively. Figure 5 earlier shows that the average yield of green bonds is more rewarding than that of conventional bonds throughout the whole observed periods for both the US and Europe, and the positive impact of green bonds found from Table 6 are in line with the results. Meanwhile, Figure 5 shows that green bonds in China have a higher payoff in the first half of observed periods and then outperformed by conventional bonds in the second half of the observed periods. This again is in line with the regression result of the impact of the green bond being -0.137. Comparably, the Hausman Taylor model gives quite different results. The green bonds label is insignificant for China, the US and Europe. This can be interpreted as the green bond label having no impact on the yield performance. However, the insignificant impact of the green bond label fails to decode the notable yield differences from Figure 5.

In addition to it, the impact of yield spread for all models is positive and significant for China, the US and Europe. On top of that, yield spread brings the greatest influence on mid yield with great magnitude. It is also worthwhile noticing that the results differ greatly within pooled OLS and the Hausman Taylor model. Opposite to the great difference in pooled OLS and the Hausman Taylor model shown in yield spread, the impacts of bid yield change are consistent within the two models. For China, pooled OLS has a significant impact of 1.178 and the Hausman Taylor model gives 1.065. Bid yield change shows a higher influence on the bonds for the US with the impact being 3.708 and 3.287 from the two models significantly. As for Europe, both models give an insignificant impact on bid yield change. Maturity shows a generally positive influence from all models. However, the results from pooled OLS are more significant than the results from the Hausman Taylor model. For China, the influence of maturity from OLS is 0.282 and is insignificant from the Hausman Taylor model. For the US, the impact is around 0.06 for both models and around 0.5 for Europe.

Generally speaking, results from pooled OLS show almost significant. The impact of the green bond label and yield spread show the most significant difference between OLS and the Hausman Taylor model. The green bond label from the Hausman Taylor shows all insignificant, and in pooled OLS, it is negative for China and positive for the US and Europe. And the yield spread differs greatly within the two models for all observed countries. For the rest variables, the results are consistent within the two models. Bid yield change gives generally positive influence over bonds' yield for China and the US whereas it has no impact on the bonds in Europe. Additionally, maturity positively affects bonds' yield albeit with a comparably small magnitude.

Lastly, to compare with the information exposed from Figure 5, pooled OLS does a better job in capturing the green bond label impact. The regression results of negative greenium for China and positive greenium for the US and Europe align well with the visualized findings. Conversely, the Hausman Taylor model gives insignificant coefficients for the green bond label and thus fails to capture the signal.

5.2 Hypothesis 2: Incorporation with Macro variables

Hypothetically, the worldwide pandemic caused by Coronavirus has traumatised the economy of all countries. Henceforth, the second research question is aiming to discover if an economic recession will bring the bond yield down. To do so, macro variable GDP is incorporated.

	China		US		Europe	
VARIABLES	Pooled OLS	HT	Pooled OLS	HT	Pooled OLS	HT
Yld Spread	45.03***	109.7***	10.67***	4.887***	5.926***	0.757***
(TVendogenous)	(0.224)	(3.318)	(0.112)	(0.501)	(0.140)	(0.261)
Bid Yld Change	0.559*	0.326***	0.708***	0.736***	9.14e-05	6.19e-05
(TVexogenous)	(0.299)	(0.0905)	(0.212)	(0.133)	(0.000295)	(0.000143)
Green Bond	-0.141***	-62.83	0.167***	-3.760	-0.0303***	0.530
(TIendogenous)	(0.0184)	(133.8)	(0.00866)	(8.127)	(0.00482)	(0.557)
Maturity	0.288***	-1.096	0.0548***	0.0531*	0.0652***	0.0181
(TIexogenous)	(0.00640)	(4.378)	(0.000271)	(0.0291)	(0.00129)	(0.0462)
GDP	-0.0773***	-0.0824***	0.114***	0.117***	-0.0351***	0.0914***
(TVexogeneous)	(0.00681)	(0.00208)	(0.00458)	(0.00288)	(0.00187)	(0.0102)
Constant	0.486***	33.33	-1.232***	1.116	-0.597***	-0.691***
	(0.119)	(78.02)	(0.0804)	(4.180)	(0.0136)	(0.250)
Observations	15,360	15,360	7,380	7,380	7,700	7,700
Number of Bonds	30	30	18	18	20	20
R-squared	0.764		0.866		0.293	

***, ** Statistically significant at the 1% level and 5% level, respectively.

Table 7: Empirical results for Pooled OLS and the Hausman Taylor methods for incorporating GDP. (OLS: Mid $\text{Yld}_{it} = \beta_0 + \beta_1 \text{BidyChg}_{it} + \beta_2 \text{Yld spread}_{it} + \beta_3 \text{GDP}_{it} + \delta_1 \text{GB}_i + \delta_2 \text{Maturity}_i + \epsilon_{it}$, HT: Mid $\text{Yld}_{it} = \beta_0 + \beta_1 \text{BidyChg}_{it} + \beta_2 \text{Yld spread}_{it} + \beta_3 \text{GDP}_{it} + \delta_1 \text{GB}_i + \delta_2 \text{Maturity}_i + u_i + \epsilon_{it}$, with $[x_1] = [\text{BidyChg}_{it}, \text{GDP}_{it}]$, $[x_2] = [\text{Yld spread}_{it}]$, $[z_1] = [\text{Maturity}_i]$ and $[z_2] = [\text{GB}_i]$.) The results are presented for China, the US and Europe respectively. Due to the incorporation of the GDP factor for 2020 and 2021, bonds that were issued in 2019 are disregarded in this case. The periods of the database for China, the US and Europe are 2020-08-07 - 2021-12-31, 2020-11-16 - 2021-12-31 and 2020-12-11 - 2021-12-31 with 30, 18 and 20 bonds respectively. It is worthwhile noticing that it does not make difference when GDP is treated as time-varying endogenous. The results are presented in Table 19 in section 8 for interested readers.

First and foremost, results for most variables do not differ much compared with the baseline model Table 6. To be more specific, bid yield change stays still positive for China and the US for both OLS and the Hausman Taylor model and insignificant for Europe. The only notable change is the impact decreases for all countries. For example, the influence drops from around 1 to 0.5 for China and 3.5 to 0.7 for the US. The green bond label from pooled OLS model gives consistent and similar results Table 6 for China and the US. That is, negative greenium as -0.141(-0.137 from Table 6) for China and positive greenium as 0.167(0.119 from Table 6) for the US. However, the results show oppositely for Europe where the impact of the green bond label is negative as -0.0303(0.235 from Table 6). This briefly tells that the impact green bond label is not affected much when the macro variable is added for OLS regression. Meanwhile, the Hausman Taylor model for all countries still presents an insignificant green bond label coefficient. In terms of maturity, OLS models stay consistent from Table 6 to Table 7 for China and the US with values being 0.288(0.282 from baseline model) and 0.0548(0.0580 from baseline model). The exception shows for Europe where the value changes from 0.208 to 0.0652, so less impact. From the perspective of the Hausman-Taylor model, the coefficient of maturity for China and Europe still remains insignificant. Only the coefficient for the US is significant at 0.0531.

The vision is then turned to GDP, namely, the macro variable. It is worth noticing that all results show the significance of both models for all countries. For China, GDP has a negative impact of around -0.8 for both models. This indicates during the pandemic period, one unit growth in GDP will bring down the average bond by -0.08. For the US, the growth in the US helps to energize bond yield with around 0.115

bps for both models. Moreover, for Europe, OLS gives a negative value of -0.0351 and a positive value from the Hausman Taylor model of 0.0914. Conclusively, pooled OLS indicate that growth in GDP brings down bond yields for China and Europe but increases that in the US. The results from the Hausman Taylor model, however, suggest that GDP growth has a positive impact on yield returns for the US and Europe but negative for China. Overall, after adding GDP as a macro factor, the results for the green bond label do not differ much for both pooled OLS and the Hausman Taylor model. GDP enters with a negative impact on the bond yield of China.

Silva and Cortez [2016] have devoted their research specifically to the investigation of green bonds' performance during crisis times and non-crisis times. They find that green bonds are more resilient during crisis time compared to non-crisis time, where the bonds perform better in the former period. Moreover, green bonds in the US outperform other conventional socially responsible bonds during difficult times. To link green bonds and economic situation directly, in other words, to differentiate the performance of green bonds and conventional bonds specifically under a pandemic environment, an interaction term is imposed in the pooled OLS regression.

	pooled OLS		
VARIABLES	China	US	Europe
Yld Spread	45.03***	10.67***	5.624***
	(0.224)	(0.112)	(0.138)
Bid Yld Change	0.556*	0.709***	0.000139
	(0.299)	(0.212)	(0.000288)
Green bond	0.584**	0.299*	0.172***
	(0.230)	(0.159)	(0.0119)
Maturity	0.288***	0.0548***	0.0637***
	(0.00640)	(0.000271)	(0.00127)
GDP	-0.0558***	0.118***	-0.00116
	(0.00962)	(0.00647)	(0.00260)
GDP*GB	-0.0430***	-0.00760	-0.0662***
	(0.0136)	(0.00915)	(0.00359)
Constant	0.124	-1.298***	-0.675***
	(0.165)	(0.113)	(0.0140)
Observations	15,360	7,380	7,700
Number of Bonds	30	18	20
R-squared	0.764	0.866	0.323

***, ** Statistically significant at the 1% level and 5% level, respectively.

Table 8: Pooled OLS regression constructed based on subsubsection 4.1.2 with an interaction term. OLS: $\text{Mid Yld}_{it} = \beta_0 + \beta_1 \text{BidyChg}_{it} + \beta_2 \text{Yld spread}_{it} + \beta_3 \text{GDP}_{it} + \beta_4 \text{GDP}_{it} * \text{GB}_i + \delta_1 \text{GB}_i + \delta_2 \text{Maturity}_i + \epsilon_{it}$.

Zoom in to the impact of the interaction term, it shows a negative influence on China and Europe with the value being -0.0430 and -0.0662 significantly. This indicates that green bond does differentiate from conventional bonds with the incorporation of economics. Moreover, it can be interpreted as with a unit increase in GDP, the yields of green bonds are comparably less than conventional bonds in the pandemic period from 2019 to 2020. This result is in line with Silva and Cortez [2016] where they find that green bonds still underperform conventional bonds during crises, however, the performance improves than in non-crisis times. In other words, yields of green bonds will increase when GDP declines. Referring to the information exposed to China from Figure 5, the green bonds are more rewarding than conventional bonds in the first half of the pandemic period and the table turns in the second half period. This can be interpreted that the economy was suffering from the pandemic at an early stage which is also the resilient phase for green bonds. With the control policy imposed and the situation of Covid getting better, the economy revives later on and green bonds are surpassed by conventional bonds again. The same story happens also in Europe as one can observe from Figure 5, in the first half of period green bonds outperform conventional bonds with significant differences. With time moving on, the gap shrinks and conventional bonds almost catch up with green bonds in a way they have the same yields. At the same time, the interaction term for the US shows an insignificant impact. This indicates that green bond does not behave as anything distinguishable from conventional bonds even considering the economic situation. This again is in line with Figure 5 where the green bonds are always more rewarding than conventional bonds throughout the whole crisis period and the gap in-between is consistent and stable.

Summarily, green bonds do behave differently in different stages of economics. For China and Europe, green bonds are more resilient during volatile periods and the advantage of resilience fades away when the economy grows back. As for the US, the green bonds are not affected by GDP within the whole observed pandemic period. In other words, even the reviving of the economy does not defeat the outperformance of the green bonds.

5.3 Hypothesis 3: Incorporation with Fama French size premium

	China			US			Europe		
VARIABLES	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Yld Spread	39.81***	18.35***	39.81***	11.27***	39.94***	11.26***	20.78***	40.83***	20.82***
(TVendogenous)	(0.197)	(0.409)	(0.197)	(0.452)	(1.555)	(0.452)	(0.225)	(0.318)	(0.225)
Bid Yld Change	0.994***	0.952***	0.993***	4.042***	3.670***	4.039***	-0.00126	0.000178	-0.00123
(TVexogenous)	(0.254)	(0.126)	(0.254)	(0.337)	(0.260)	(0.337)	(0.00172)	(0.000707)	(0.00172)
Green Bond	-0.315***	27.31	-0.425*	0.0600***	9.286	0.611	0.233***	38.24	0.00666
(TIendogenous)	(0.0169)	(81.67)	(0.230)	(0.0176)	(16.53)	(0.580)	(0.0223)	(142.0)	(0.0812)
Maturity	0.142***	-0.317	0.142***	0.0594***	0.0651	0.0594***	0.206***	-1.995	0.206***
(TIexogenous)	(0.00619)	(2.678)	(0.00619)	(0.000520)	(0.0618)	(0.000520)	(0.00568)	(8.905)	(0.00568)
Size premium	-1.045***	-0.517***	-1.075***	1.512***	1.362***	1.581***	0.240***	0.206***	0.220***
TVexogenous	(0.0637)	(0.0330)	(0.0899)	(0.0724)	(0.0563)	(0.102)	(0.00691)	(0.00289)	(0.00979)
Size premium*GB			0.0606			-0.137			0.0401***
(Interaction term)			(0.127)			(0.144)			(0.0138)
Constant	2.003***	-9.461	2.058***	-5.036***	-10.94	-5.311***	-3.496***	-10.23	-3.387***
	(0.118)	(34.57)	(0.165)	(0.292)	(8.672)	(0.411)	(0.0563)	(21.62)	(0.0677)
Observations		14952			5,340			10,392	
Number of Bonds		24			20			24	

***, ** Statistically significant at the 1% level and 5% level, respectively.

Table 9: Empirical results with size premium extended as an explanatory variable. Model (1) and (2) are the practice for subsubsection 4.1.3 for pooled OLS and Hausman-Taylor model respectively. (OLS: $\text{Mid Yld}_{it} = \beta_0+\beta_1\text{BidyChg}_{it}+\beta_2\text{Yld spread}_{it}+\beta_3\text{SMB}_t+\delta_1\text{GB}_i+\delta_2\text{Maturity}_i+\epsilon_{it}$, HT: $\text{Mid Yld}_{it} = \beta_0+\beta_1\text{BidyChg}_{it}+\beta_2\text{Yld spread}_{it}+\beta_3\text{SMB}_t+\delta_1\text{GB}_i+\delta_2\text{Maturity}_i+u_i+\epsilon_{it}$, with $[x_1] = [\text{BidyChg}_{it}, \text{SMB}_t]$, $[x_2] = [\text{Yld spread}_{it}]$, $[z_1] = [\text{Maturity}_i]$ and $[z_2] = [\text{GB}_i]$). Model (3) is the pooled OLS regression incorporating an interaction term (Size premium $*$ GB). The model presented as: $\text{Mid Yld}_{it} = \beta_0 + \beta_1\text{BidyChg}_{it} + \beta_2\text{Yld spread}_{it} + \beta_3\text{SMB}_t + \beta_4\text{SMB}_{it} * \text{GB}_i + \delta_1\text{GB}_i + \delta_2\text{Maturity}_i + \epsilon_{it}$

Firstly, for China, size premium has a negative and significant impact with a value being -1.045 from

Model (1). This indicates that small-cap firms give out lower yields compared to big-cap firms. This finding surprisingly contradicted with Fama French 3 model hypothesis where the model argues that small-cap firms have more rewarding yields. On the other hand, Model (1) for the US and Europe shows that size premium act positively on yields. This infers that small-cap firms give out higher yields than big-cap firms. Linked the findings to Figure 7 where all countries show that small-cap firms have significantly higher yields than bigger ones. Therefore, the findings from China are anomalous. Move along to Model (2), which is also referred to as the Hausman Taylor model. One can observe that the impact of size premium inherits generally the same but with less magnitude: -1.045 to -0.517 for China, 1.512 to 1.362 for the US and 0.240 to 0.206 for Europe. This indicates both pooled OLS and the Hausman Taylor model evidence that the small firm size has a negative impact on Chinese bonds but a positive influence on the bonds in the US and Europe. The green bond label impact of all models follow generally the same patterns from the earlier findings. To be more specific, China has negative greenium for pooled OLS whereas the US and Europe have positive greenium. And green bond label shows again insignificantly for all Hausman Taylor models.

Model (3) takes an interaction term that combines green bond label and size premium into account. The interaction term shows insignificantly for both China and the US, this tells that green bonds are indifferent to the size of firms specifically compared to conventional bonds. In other words, green bonds and conventional bonds react equivalently to the size factor. However, the case fails for Europe, where the interaction term is positive and significant with a value being 0.0401. Comprehensively, this signals that green bonds have a greater reaction to the size factor with an extra 0.0401 compared to conventional bonds.

The regressions tell that size premium plays an active role in yields. In China, big firms have higher yields and the opposite case holds for both the US and Europe. Meanwhile, green bonds in Europe respond more positively than conventional bonds and they respond equivalently to conventional bonds for China and the US.

5.4 Combine with all database

Lastly, one large data set that combines all countries are generated. To be more specific, all bonds data are gathered together based on the shared existing period from 2021-07-29 to 2022-04-21. The attributes consist of only yield spread, bid yield change, the green label and maturity. The baseline model subsubsection 4.1.1 is examined with this large database to gather a broader view of the impact of the green bond label. Table 10 indicates that pooled OLS and the Hausman Taylor model differ from each other in terms of green bond and maturity. Regarding the results from OLS, the green bond has a positive and significant impact on yield return with 0.154 bps. This stands that if the bond is green labelled then it brings an extra 0.154 yields. Meanwhile, the Hausman Taylor model gives comparably different results where the green label does not bring any difference. The same finding holds for maturity where OLS tells that maturity gives a positive impact with 0.0586 and the Hausman Taylor model shows an insignificant impact of maturity. Besides the above findings, yield spread and bid yield change show similar findings from both OLS and Hausman Taylor model where the latter is strongly positive of being around 23 and the latter is insignificant.

VARIABLES	OLS	HT
Yld Spread	23.67***	23.71***
(TVendogenous)	(0.250)	(0.352)
Bid Yld Change	0.000235	0.00131
(TVexogenous)	(0.00587)	(0.00197)
Green Bond	0.154***	0.216
(TIendogenous)	(0.0225)	(0.669)
Maturity	0.0586***	-0.0549
(TIexogenous)	(0.000946)	(0.886)
Constant	0.290***	1.459
	(0.0260)	(9.151)
Observations	21,360	
Number of id	80	

***, ** Statistically significant at the 1% level and 5% level, respectively.

Table 10: Empirical results for all data combined together including China, the US and Europe. (OLS: $\text{Mid Yld}_{it} = \beta_0 + \beta_1 \text{BidyChg}_{it} + \beta_2 \text{Yld spread}_{it} + \delta_1 \text{GB}_i + \delta_2 \text{Maturity}_i + \epsilon_{it}$.) The starting date depends on the latest bond issuing date and the observation period is: 2021-07-29 to 2022-04-21.

6 Robustness Check

This section validates the plausibility from several perspectives. First, a multicollinearity check is implemented by variance inflation factor to investigate if there is a direct correlation within time-varying variables, that is, the potential correlation between the yield spread and bid yield change. The results show that $\text{VIF}_{yldspread} = 1$ and $\text{VIF}_{BidYChg} = 1$. Therefore, the yield spread is orthogonal to bid yield change and thus has no linear correlation. In other words, it's safe to incorporate both yield spread and bid yield change in the same model. The following section subsection 6.1 argues the potential difference in results brought by different treatments to the green bond label, size premium and the GDP factor. Furthermore, the plausibility of the results of the Hausman Taylor model is examined with the yield curve in subsection 6.2. Finally, in subsection 6.3 backtesting is experimented with by firstly simulating a series of mid yield data and comparing the predictions of both pooled OLS and the Hausman Taylor model.

6.1 Results of twisting X and Z for the Hausman Taylor

In subsubsection 4.3.2, Table 4 twists the hypotheses by applying the green label as exogenous. Table 11 displays the results for the renovated hypotheses. By observing mainly the difference between before twisting and after twisting, it is evident that only insignificant values such as green bond and maturity have changed. Meanwhile, yield spread, bid yield change, GDP and size premium persist at the exact same values. Therefore, the conclusions can be safely drawn that the Hausman-Taylor model is sensitive to the shift of the green bond label from endogenous and exogenous. However, with values being insignificant, the change in the results does not bring any effects on the findings.

Table 5 introduced the second twisting experiment where the GDP factor and the size factor are treated as exogenous. Table 19 and Table 20 in Appendix state the empirical results for such analysis. The comparisons show identical results between GDP and size premium being treated as either exogenous or endogenous.

Therefore, one can infer that the Hausman Taylor model is not sensitive to the control of GDP and size factor. Gathering the above information together, the Hausman Taylor model is only sensitive with the switch of the green bond label being endogenous to exogenous, however, the different results are barely informative since the values are insignificant.

6.2 Yield Curve Check

The empirical results Table 6, Table 7 and Table 9 from 3 hypotheses suggest that yield increases with maturity for pooled OLS whereas the Hausman Taylor model indicates there is no relationship in-between. This connection can be verified by the yield curve. If the results from one model present consistent insights into the yield curve, it, therefore, hints that the results are robust.

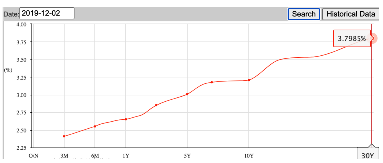
The yield curve demonstrates the explicit relationship between maturity and yield. The yield curve is constructed by connecting the yields of bonds from different maturity dates but with the same ranking. Normally yield curve has an upward increasing pattern, that infers longer maturity bonds have higher yields than that shorter maturity bonds. There are two major underlying reasons. First, bonds with long maturity are more sensitive and are more affected by the movement of interest rates. When the interest rate increases, the yield or coupons of the bond is less attractive and therefore bring down the value or the price of the bond. And the bond value increases when the interest rate goes down the other way around. Comparably, bonds with shorter maturity are less affected by the interest rate and therefore less volatile. Secondly, longer maturity means a higher risk of bearing the issuer's credit defaults or bankruptcy. In other words, it is difficult to predict the future of a company and also the likelihood of the existence of a corporate. Therefore, to compensate for the volatility and the risk of uncertainty, the yields of longer-maturity bonds are generally higher than short-maturity bonds.

It is also interesting to note that the inverted yield curve signals a recession where shorter-term yields are higher than longer-term yields. It can be explained when there is a recession, investors prefer to invest in long-term bonds with the belief they are safer. Therefore, the demands for long-term bonds increase and the issuers tend to lower the yields since they have no stress of seeking buyers.

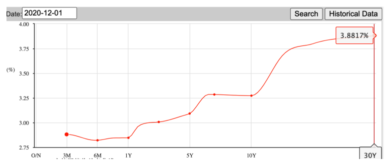
Figure 8 demonstrates that the yield curves display upward trends from 2019 to 2021 for China, the US and Euro areas. In addition, the yield curve is more volatile in the US in 2019 where the curve shows an inverted bell shape. Generally, Figure 8 reveals that long-term yields are higher than short-term yields, thus the yields return increases with maturity extension. To link to the regression results from earlier sections, only pooled OLS is able to give a positive and significant impact of maturity on yields consistently, whereas the impact of maturity shows mostly insignificant from the Hausman Taylor model. Therefore, the robustness check suggests that the Hausman Taylor model is unable to expose the yield curve while pooled OLS conveys the information in the results.

6.3 Backtesting

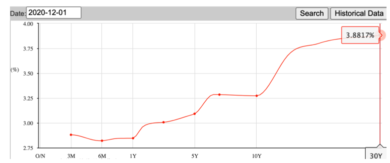
Lastly, the comparison of the prediction power between the Hausman Taylor model and pooled OLS is imposed to discover if these two models truly can reproduce the data and also help to dig into which one is more convincing. To achieve the goal, a series of simulated data endowed with the same characteristics as the realized data are needed. This is where backtesting comes in. The idea of backtesting is to use the past data and treat them as realized, then fit them to a model to study the data and predict the future values. Subsequently, the predictions are compared with the realized data to examine how well the models



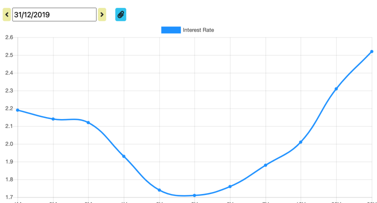
(a) 2019 China Government Bond yield curve



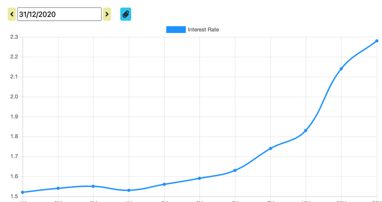
(b) 2020 China Government Bond yield curve



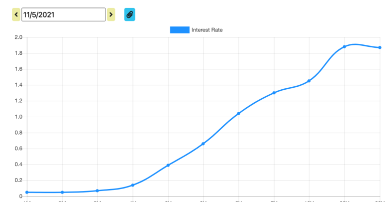
(c) 2021 China Government Bond yield curve



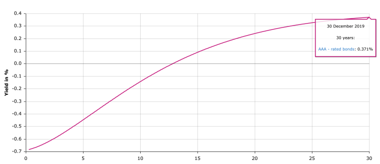
(d) 2019 US Treasuries yield curve



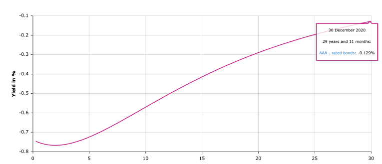
(e) 2020 US Treasuries yield curve



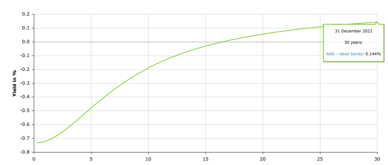
(f) 2021 US Treasuries yield curve



(g) 2019 Europe euro area AAA rated yield curve



(h) 2020 Europe euro area AAA-rated yield curve



(i) 2021 Europe euro area AAA-rated yield curve

Figure 8: Yield curve from 2019 to 2021 for China, the US and Europe with euro areas only. Data source: **China Government Bond Yield Curve** [a], **US Treasuries Yield Curve**[b] and **Euro area Yield Curve** [c].

[a]https://www.climatebonds.net/market/best-practice-guidelines
[b]https://www.ustreasuryyieldcurve.com/
[c]https://www.ecb.europa.eu/stats

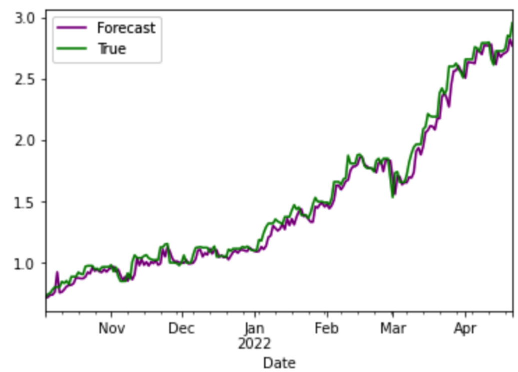
	Green bond label as endogenous								
	China			US			Europe		
VARIABLES	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Yld Spread	16.40***	109.7***	18.35***	39.70	4.887***	39.94**	45.12***	0.757***	40.83***
(TVendogenous)	(0.391)	(3.318)	(0.409)	(1.285)	(0.501)	(1.555)	(0.380)	(0.261)	(0.318)
Bid Yld Change	1.065***	0.326***	0.952***	3.287***	0.736***	3.670	0.000681	6.19e-05	0.000178
(TVexogenous)	(0.126)	(0.0905)	(0.126)	(0.251)	(0.133)	(0.260)	(0.000862)	(0.000143)	(0.000707)
Green Bond	19.35	-62.83	27.31	-3.448	-3.760	9.286	-2.428223	0.530	38.24
(TIendogenous)	(35.93)	(133.8)	(81.67)	(4.941)	(8.127)	(16.53)	(3.6942)	(0.557)	(142.0)
Maturity	0.749	-1.096	-0.317	0.0729**	0.0531*	0.0651	0.5272*	0.0181	-1.995
(TIexogenous)	(1.329)	(4.378)	(2.678)	(0.0292)	(0.0291)	(0.0618)	(0.307)	(0.0462)	(8.905)
GDP		-0.0824***			0.117***			0.0914***	
(TVexogenous)		(0.00208)			(0.00288)			(0.0102)	
Size premium			-0.517***			1.362***			0.206***
(TIexogenous)			(0.0330)			(0.0563)			(0.00289)
Constant	-10.51	33.33	-9.461	0.686	1.116	-10.94	-4.3574*	-0.691***	-10.23
	(21.15)	(78.02)	(34.57)	(2.488)	(4.180)	(8.672)	(1.4090)	(0.250)	(21.62)
	Green bond label as exogenous								
Yld Spread	16.40***	109.7***	18.35***	39.70	4.887***	39.94**	45.12***	0.757***	40.83***
(TVendogenous)	(0.391)	(3.318)	(0.409)	(1.285)	(0.501)	(1.555)	(0.380)	(0.261)	(0.318)
Bid Yld Change	1.065***	0.326***	0.952***	3.287***	0.736***	3.670	0.000681	6.19e-05	0.000178
(TVexogenous)	(0.126)	(0.0905)	(0.126)	(0.251)	(0.133)	(0.260)	(0.000862)	(0.000143)	(0.000707)
Green Bond	**0.519**	**-1.612**	**-0.746**	**0.386**	**0.0996**	**5.881**	-2.428223	**0.0460**	**-1.620**
(TIexgenous)	(2.126)	(4.306)	(2.046)	(4.941)	(1.146)	(311.4)	(3.287)	(0.121)	(2.359)
Maturity	**4.145**	**-7.750**	**3.604**	**0.0273**	**-0.0969**	**8.181**	**-1.110**	**0.0107**	**2.286**
(TIendogenous)	(4.005)	(6.220)	(3.507)	(0.0375)	(0.387)	(436.2)	(3.007)	(0.0462)	(1.673)
GDP		-0.0824***			0.117***			0.0914***	
(TIVexogenous)		(0.00208)			(0.00288)			(0.0102)	
Size premium			-0.517***			1.362***			0.206***
(TIexogenous)			(0.0330)			(0.0563)			(0.00289)
Constant	**-16.04**	32.00	2.058***	-0.232	2.994	-192.1	3.470	-0.416	-16.34*
	(17.95)	(27.93)	(0.165)	(0.895)	(9.891)	(10,032)	(16.82)	(0.271)	(9.415)
Observations	18,690	15,360	14,952	6,942	7,380	5,340	10,392	7,700	10,392
Number of Bonds	30	30	24	26	18	20	24	20	24

***, ** Statistically significant at the 1% level and 5% level, respectively.

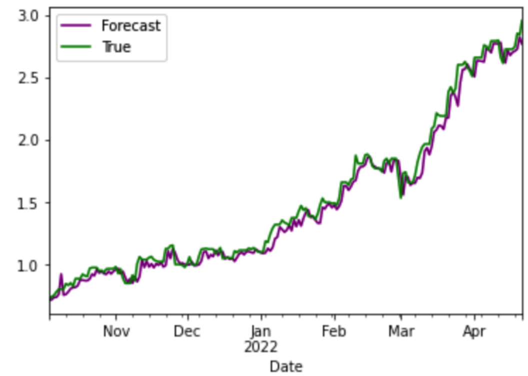
Table 11: Comparison results from the Hausman Taylor model for treating green bond labels as endogenous and exogenous. Model (1) stands for the first hypothesis which is also referred to as the baseline model. Model (2) stands for the GDP factor extended model. Model (3) stands for the third hypothesis where size premium is incorporated. The upper part of the table presents the earlier results that apply the green bond label as endogenous. The below part of the table presents the results that treat green bonds as exogenous. Distinct values are marked in bold.

are fitted. In this case, the Garch model is employed as the fitted model. Figure 9 shows the forecast of mid yield produced by the Garch model for China, the US and Europe for respective certain observed periods. It is evident the forecasts are highly aligned with realized data and are able to capture well the ups and downs movements. Henceforth, one is conformable to utilize the data simulated by the Garch model and treat them as realized data for the further practice for pooled OLS and the Hausman Taylor model. By applying the regression coefficients from pooled OLS and the Hausman Taylor model from Table 6, the predictions of both models are presented and compared in Figure 10.

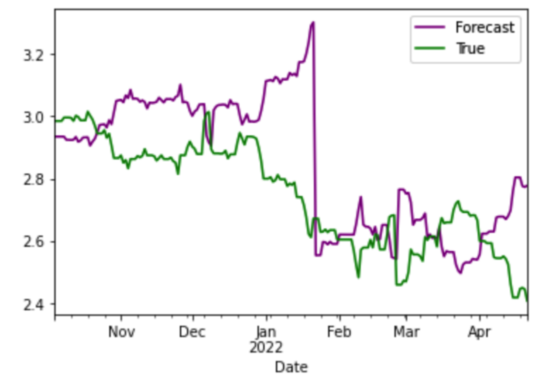
Zoom out and observe from a macro view: the pink dots are more congregating to the blue line and the yellow dots are more spreading out and scattered. This accordingly tells that pooled OLS gives comparably more plausible predictions than the Hausman Taylor model. For China, OLS predictions do a better job at an early stage where the predictions are pretty concentrated around the Garch forecast for ID 3 and ID 5. The predictions drift apart at a later stage. However, it turns oppositely for ID 21, where the predictions perform remarkably at the second half stage and the predictions meet almost with the Garch forecast. Meanwhile, the Hausman Taylor model shows a similar predictions pattern to OLS by observing the distribution of the dots, however, the predictions generally perform worse than OLS. For the US, OLS also gives decent predictions in the first half stage and the performance decay over time for ID 3 and ID 25 with the Garch forecast drifting from OLS predictions. For ID 11, the predictions of OLS are not as plausible as previous, it still outperforms the Hausman Taylor predictions with a comparably smaller gap. For Europe, it is interesting to note that both OLS and Hausman Taylor models give better performance in prediction throughout the whole period than China and the US. OLS regression still gives better predictions by presenting more congregated predictions and the ability to capture the upward trend in the second half period.



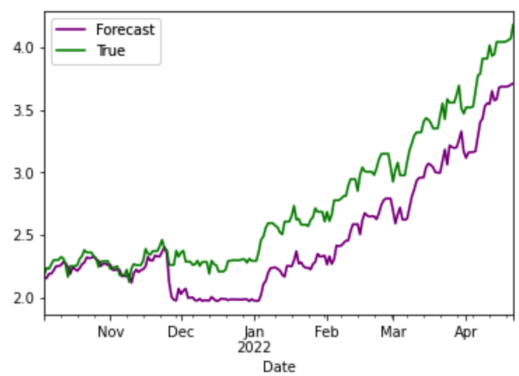
(a) China green bond ID 3, maturity of 3 years



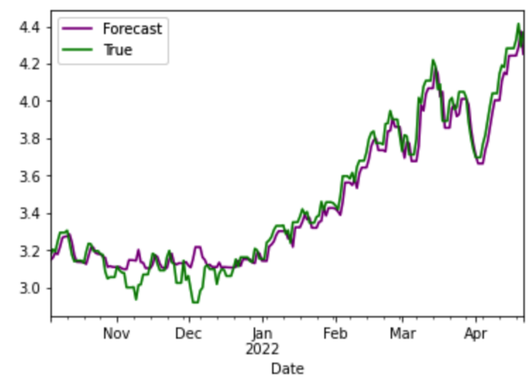
(b) China green bond ID 5, maturity of 5 years



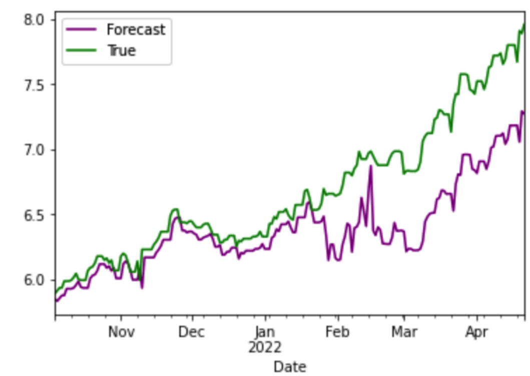
(c) China green bond ID 21, maturity of 3 years



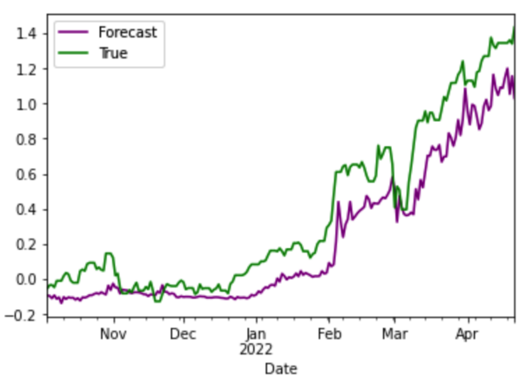
(d) USA green bond ID 3, maturity of 10 years



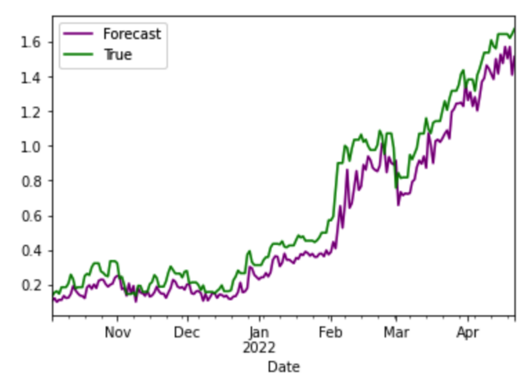
(e) USA green bond ID 25, maturity of 30 years



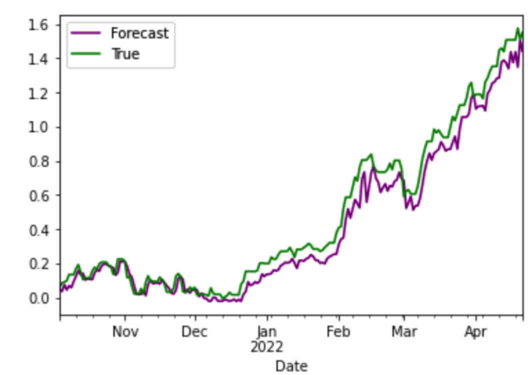
(f) USA green bond ID 11, maturity of 60 years



(g) Europe green bond ID 1, maturity of 6 years

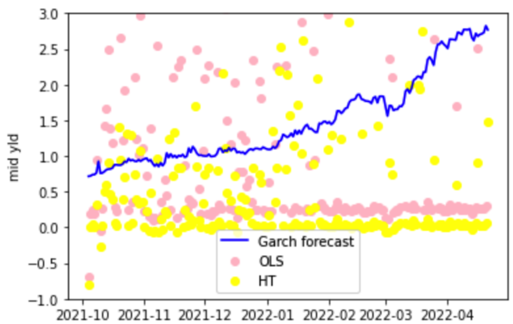


(h) Europe green bond ID 17, maturity of 7 years

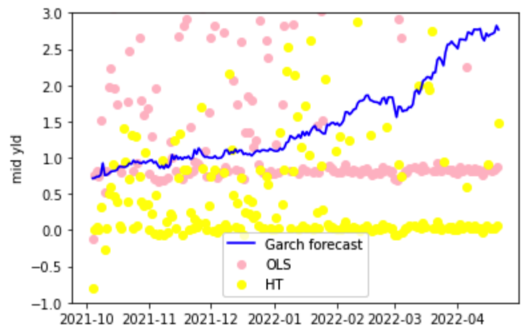


(i) Europe green bond ID 9, maturity of 9 years

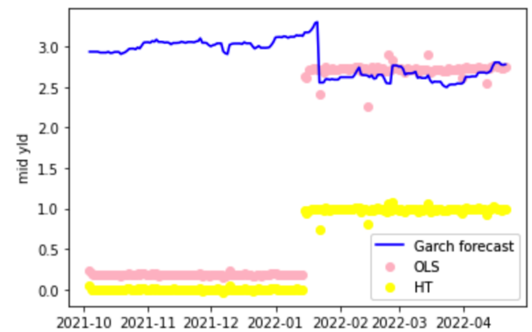
Figure 9: Comparison of Garch forecast mid yield and realized mid yield. Only random samples of bonds are selected and predicted with the main aim to see the prediction performance.



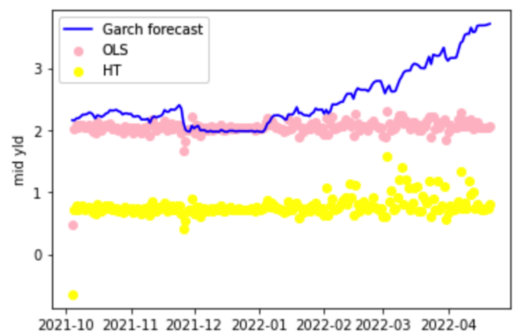
(a) China green bond ID 3, maturity of 3 years



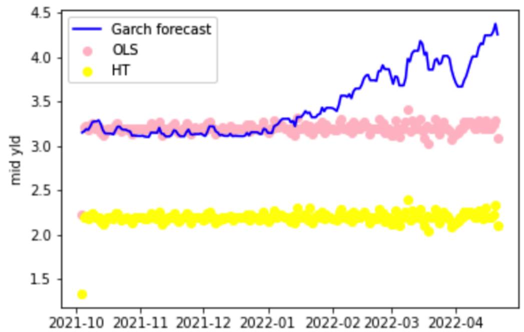
(b) China green bond ID 5, maturity of 5 years



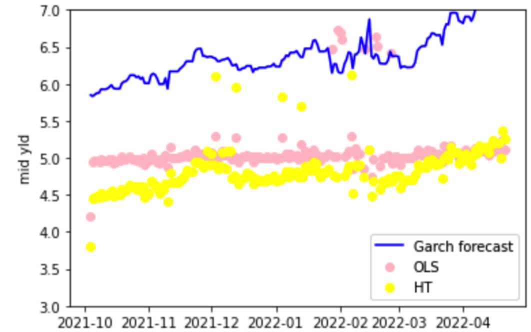
(c) China green bond ID 21, maturity of 3 years



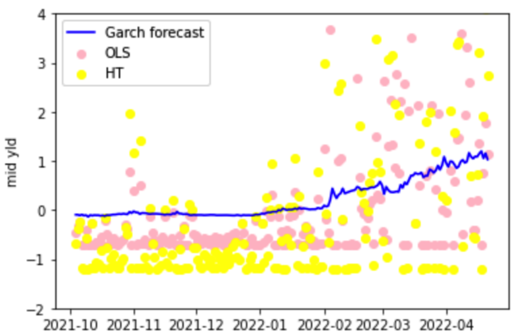
(d) USA green bond ID 3, maturity of 10 years



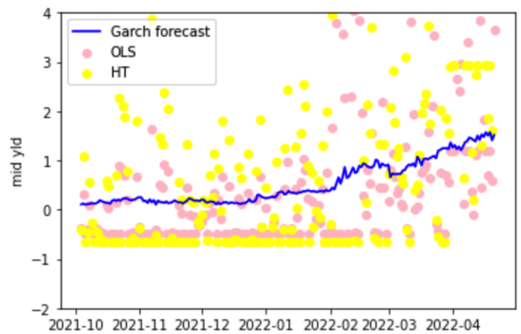
(e) USA green bond ID 25, maturity of 30 years



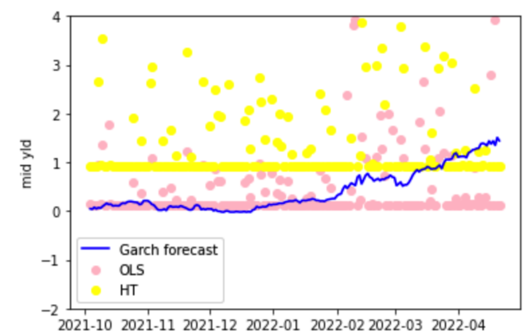
(f) USA green bond ID 11, maturity of 60 years



(g) Europe green bond ID 1, maturity of 6 years



(h) Europe green bond ID 17, maturity of 7 years



(i) Europe green bond ID 9, maturity of 9 years

Figure 10: Pooled OLS and the Hausman Taylor model prediction using the regression coefficients from Table 6, only significant results are employed. Random samples of bonds are selected for the main aim of gaining insights of prediction performance.

7 Conclusion and Discussion

In this section, the key findings, as well as contributions of this paper, are provided. Furthermore, a discussion followed with the limitation and potential for further research.

7.1 Conclusion

This research extends from Zerbib [2019] in a way to explore the most recent evolution of green bonds. The collected bonds are issued between 2019 and 2021 in China, the US and Europe. A complete balanced database is constructed based on the same matching scheme from Zerbib [2019] where a pair of conventional bonds and green bonds from the same issuer and with a similar issuing date are selected. The paper also sheds light on exploring the green bond label impact directly, which is limited by previous research applications. To also account for time-invariant heterogeneity, a special econometric model the Hausman Taylor model is also implemented for further exploration. Additionally, not many econometric models have been experimented with for green bonds research, this paper thus contributes in the way to validating the plausibility of applying the Hausman Taylor model for this topic.

The regression results from the baseline method OLS are distinct from the Hausman Taylor model. As argued earlier, the biggest challenge in using the Hausman Taylor model is to differentiate which time-varying variables are correlated with time-invariant variables. subsection 6.1 finds out the Hausman Taylor model is only sensitive to the green bond label and maturity. By treating the green bond label as exogenous and maturity as endogenous, only the impact of the two variables has changed. However, the values are still insignificant even with different treatments. This finding suggests that different treatments of the green bond label do not bring any distortions to the Hausman Taylor model. Moreover, from subsection 6.2, only pooled OLS is able to expose the agreeing information from the yield curves whereas the Hausman Taylor model fails to provide any information about the relationship between maturity and yields. Finally, backtesting is implemented for comparing the prediction power from these two models. The Garch model is applied to simulate alike data and linear predictions from both OLS and the Hausman Taylor model are imposed over the simulated data for comparison. subsection 6.3 shows that OLS performs comparably better than the Hausman Taylor model and gives a more plausible prediction, that is, more aligned with the Garch forecast. In summary, the Hausman-Taylor model is only sensitive to the different treatments of the green bond label and maturity. However, both coefficients display insignificantly with different treatments and thus have no influence on the findings. This further reveals that time-invariant heterogeneity does not disturb green bond analysis. That is, potential correlations between time-invariant and time-varying variables such as yield spread and green bond label or rating do not affect the analysis of the impact of the green bond label. This concludes that more interpretation should be imposed on pooled OLS results.

To answer the first research question, one firstly focuses on the green bond label. The most interesting finding is that the green bond label for the US and Europe are positive with 0.119 and 0.235 from the baseline model. This illustrates that green bonds issued from 2019 to 2021 are more rewarding than conventional bonds in the US and Europe. The results show the oppositely of Zerbib [2019] where the author finds negative greenium back in 2019 with the previously issued bond database. However, the finding is consistent with the green bond dynamics in China whose impact of the green label is negative at -0.137. In short, green bonds are more rewarding in the US and Europe but an exception in China.

The second research question imposes stress on the GDP factor. The interaction term that combines the green bond label and GDP exposes the direct link in-between. The results show that green bond yields decrease when GDP increases in China and Europe in the last 3 years. One should note that it is a special observed period with economic shock brought by the Covid pandemic and lockdown, the findings indicate that green bonds in China and Europe perform better when there is an economic low. This is consistent with Nofsinger and Varma [2014] and Silva and Cortez [2016] where they indicate that ESG mutual funds and green bonds are more resilient during economic crises. Meanwhile, the interaction term shows insignificantly for the US which infers that the green bonds are more robust and do not swing with economic shifts.

The third research question is followed up with size premium. Here, an interaction term that combines size premium and the green bond label is again employed to reveal the relationship between these two factors. The results present that green bonds do not benefit independently from small firms. However, the case is not dismissed for Europe. The interaction terms show positively with 0.0401 and it indicates that small firms give an extra 0.0401 bps higher yields for green bonds.

Lastly, all the data are combined together for the baseline model test. The results show a positive green bond label impact with 0.154. This tells that green bonds are promising from a more universal perspective. What's more, maturity has a positive impact on the mid yield which is in line with the yield curve.

The above findings present interesting results that green bonds are turning out to be more rewarding than conventional bonds nowadays. This is striking and exciting since the investors can make ethical investment choices while not losing profits at all.

7.2 Limitation and further research

The paper employs only 15 corporates from China, the US and Europe respectively, which leaves to around 30 bonds for each country and region. This is caused by the matching scheme where a pair of a green bond and a conventional bond has to coexist. Therefore, the database is limited to some extent and the analysis is restricted to present a thorough insight. For future research, larger database can be employed and thus provide more comprehensive results.

For green bonds analysis, the most common methodology is to apply fixed effects to account for time-invariant heterogeneity. In this paper, it is found that time-invariant heterogeneity does not disturb the analysis and pooled OLS regression generally gives better performance. However, there exist a gap of comparing the green bond label impact and greenium between pooled OLS and fixed effect, where green bond label is time-invariant and greenium is time-variant. Therefore, further research could be setup to explore and measure the difference.

References

António Afonso, Pedro Gomes, and Philipp Rother. What'hides' behind sovereign debt ratings? 2007.

Xiang Ao. An introduction to hausman-taylor model. *Harvard Business School*, page Research Computing Services, 2009.

Demir Bektić, Josef-Stefan Wenzler, Michael Wegener, Dirk Schiereck, and Timo Spielmann. Extending fama–french factors to corporate bond markets. *The Journal of Portfolio Management*, 45(3):141–158, 2019.

Mohamed Amine Boutabba and Yves Rannou. Investor strategies in the green bond market: The influence of liquidity risks, economic factors and clientele effects. *International Review of Financial Analysis*, 81:102071, 2022.

Darrell Duffie and Kenneth J Singleton. *Credit risk: pricing, measurement, and management.* Princeton university press, 2012.

Gianfranco Gianfrate and Mattia Peri. The green advantage: Exploring the convenience of issuing green bonds. *Journal of cleaner production*, 219:127–135, 2019.

Jerry A Hausman and William E Taylor. Panel data and unobservable individual effects. *Econometrica: Journal of the Econometric society*, pages 1377–1398, 1981.

Patrick Houweling and Jeroen Van Zundert. Factor investing in the corporate bond market. *Financial Analysts Journal*, 73(2):100–115, 2017.

Ortenca Kume and Charlie Weir. Causality relationship between bond ratings and credit spreads. *Available at SSRN 2316615*, 2013.

John Nofsinger and Abhishek Varma. Socially responsible funds and market crises. *Journal of Banking & Finance*, 48:180–193, 2014.

Florinda Silva and Maria Ceu Cortez. The performance of us and european green funds in different market conditions. *Journal of Cleaner Production*, 135:558–566, 2016.

Mohamed Ben Slimane, DD Fonseca, and Y Mahdani. Facts and fantasies about the green bond premium, 2020.

Dragon Yongjun Tang and Yupu Zhang. Do shareholders benefit from green bonds? *Journal of Corporate Finance*, 61:101427, 2020.

Olivier David Zerbib. The effect of pro-environmental preferences on bond prices: Evidence from green bonds. *Journal of Banking & Finance*, 98:39–60, 2019.

8 Appendix

8.1 Bond data statistics

Label	ID	China	the US	Europe
GB	1	CND100034297	US92343VES97	FR00140005T0
	3	CND10002R746	US898813AS93	FR00140003P3
	5	CND10002G2M9	US95040QAK04	FR0013405537
	7	CND10002C045	US744448CS82	FR0013464930
	9	CND10002GW01	US00105DAG07	FR0013514502
	11	CND10002DK40	US85513LAB09	DE000BHY0HF4
	13	CND10002F543	US906548CS94	DE000CB0HRQ9
	15	CND10002JGN6	US49374JAB98	DE000SCB0021
	17	CND10002F477	US48305QAF00	DE000DFK0GB1
	19	CND10002HXX4	US78392BAC19	ES0236463008
	21	CND100024YG4	US98422HAE62	ES0213679JR9
	23	CND100031HX9	US209111FY40	BE0974365976
	25	CND10003JF16	US040555DC57	
	27	CND10002D6F7		
	29	CND10002G5Y7		
CB	2	CND100035L30	US92343VET70	FR0013464815
	4	CND10003JGV7	US898813AR11	FR0013477296
	6	CND100034321	US95040QAH74	FR0013459815
	8	CND10004KY19	US744448CR00	FR0013398732
	10	CND10002GZQ5	US00105DAF24	FR0013477924
	12	CND100031DG3	US85513LAA26	DE000BHY0HG2
	14	CND10002K5G8	US906548CR12	DE000CB0HRT3
	16	CND100027SP0	KR6000272B37	DE000DKB0481
	18	CND10002JHK0	US48305QAG82	DE000DDA0V15
	20	CND100027NP1	US78392BAB36	XS2263652815
	22	CND100025L08	US98422HAC07	ES0413679491
	24	CND10003FN36	US209111FZ15	BE0002728096
	26	CND100036VZ4	US040555DB74	
	28	CND10002GFL8		
	30	CND10003FQB6		

Table 12: ISIN code for the selected bonds for China, the US and Europe.

Label	ID	Yld Spread Mean	Yld Spread Median	Yld Spread SD	Mid yld Mean	Mid yld Median	Mid yld SD
CB	2	0.01	0.01	0.00	0.83	0.64	0.60
	4	0.01	0.01	0.00	0.83	0.64	0.60
	6	0.01	0.01	0.00	0.83	0.64	0.60
	8	0.01	0.01	0.00	0.83	0.64	0.60
	10	0.09	0.10	0.02	5.62	6.09	0.90
	12	0.10	0.10	0.00	6.32	6.21	0.37
	14	0.10	0.10	0.00	4.23	4.08	0.64
	16	0.09	0.10	0.02	3.44	3.41	0.38
	18	0.10	0.10	0.00	7.22	7.26	0.21
	20	0.10	0.10	0.00	5.48	5.51	0.25
	22	0.01	0.01	0.00	0.83	0.64	0.60
	24	0.10	0.10	0.00	3.65	3.54	0.54
	26	0.10	0.10	0.00	4.12	4.03	0.66
	28	0.10	0.10	0.00	5.72	5.94	0.33
	30	0.09	0.10	0.02	3.91	3.82	0.47
GB	1	0.01	0.01	0.00	0.83	0.64	0.60
	3	0.01	0.01	0.00	0.83	0.64	0.60
	5	0.01	0.01	0.00	0.83	0.64	0.60
	7	0.01	0.01	0.00	0.83	0.64	0.60
	9	0.09	0.10	0.02	2.91	2.93	0.36
	11	0.10	0.10	0.00	4.99	4.98	0.51
	13	0.09	0.10	0.02	3.18	3.13	0.42
	15	0.09	0.10	0.02	3.07	3.12	0.38
	17	0.10	0.10	0.00	6.38	6.55	0.61
	19	0.10	0.10	0.00	7.30	7.43	0.43
	21	0.10	0.10	0.00	7.05	7.18	0.34
	23	0.10	0.10	0.00	3.60	3.45	0.55
	25	0.10	0.10	0.00	4.21	4.14	0.65
	27	0.10	0.10	0.00	4.91	5.03	0.58
	29	0.09	0.10	0.01	3.77	3.68	0.50

Table 13: Summary statistics of yield spread and mid yield for China selected bonds.

Label	ID	Yld Spread			Mid yld		
		Mean	Median	SD	Mean	Median	SD
CB	2	0.06	0.07	0.01	2.61	2.38	0.56
	4	0.05	0.05	0.00	3.35	3.14	0.46
	6	0.07	0.07	0.00	2.63	2.39	0.57
	10	0.05	0.05	0.00	3.17	2.97	0.45
	12	0.08	0.08	0.00	5.54	5.41	0.91
	14	0.20	0.19	0.01	3.09	3.05	0.11
	16	0.05	0.05	0.00	2.48	2.24	0.53
	18	0.10	0.10	0.00	2.50	2.43	0.44
	20	0.03	0.03	0.00	3.07	2.88	0.41
	22	0.08	0.07	0.01	2.28	2.00	0.69
	24	0.05	0.05	0.00	3.49	3.12	0.81
	26	0.05	0.05	0.00	3.48	3.28	0.40
GB	1	0.07	0.07	0.00	2.43	2.21	0.55
	3	0.05	0.05	0.00	2.56	2.31	0.59
	5	0.03	0.03	0.00	2.14	1.93	0.64
	9	0.05	0.05	0.00	3.11	2.92	0.41
	11	0.09	0.09	0.01	6.46	6.34	0.60
	13	0.15	0.14	0.01	4.11	3.83	0.52
	15	0.05	0.05	0.00	3.05	2.89	0.39
	17	0.10	0.10	0.00	1.53	1.18	0.80
	19	0.03	0.03	0.00	3.00	2.79	0.46
	21	0.05	0.05	0.00	3.09	2.86	0.59
	23	0.05	0.05	0.00	4.33	4.07	0.61
	25	0.05	0.05	0.00	3.37	3.18	0.39

Table 14: Summary statistics of yield spread and mid yield for the US selected bonds.

Label	ID	Yld Spread Mean	Yld Spread Median	Yld Spread SD	Mid yld Mean	Mid yld Median	Mid yld SD
CB	2	0.06	0.05	0.01	0.49	0.31	0.49
	4	0.05	0.05	0.00	0.18	0.01	0.39
	6	0.05	0.05	0.00	-0.03	0.00	0.24
	8	0.05	0.05	0.00	0.14	-0.04	0.34
	10	0.03	0.03	0.00	0.07	-0.10	0.41
	12	0.10	0.10	0.00	-0.09	-0.12	0.16
	14	0.07	0.07	0.00	0.60	0.44	0.41
	16	0.03	0.03	0.00	-0.04	-0.24	0.42
	18	0.05	0.05	0.00	0.99	0.86	0.22
	20	0.27	0.25	0.06	3.65	1.86	3.30
	22	0.05	0.05	0.00	-0.35	-0.38	0.06
	24	0.05	0.05	0.00	0.32	0.16	0.37
GB	1	0.07	0.07	0.00	0.14	-0.01	0.39
	3	0.05	0.05	0.00	0.40	0.20	0.49
	5	0.05	0.05	0.00	0.09	-0.03	0.35
	7	0.05	0.05	0.00	-0.02	-0.14	0.30
	9	0.03	0.03	0.00	0.20	0.05	0.40
	11	0.10	0.10	0.00	-0.18	-0.25	0.17
	13	0.06	0.07	0.01	0.43	0.31	0.38
	15	0.03	0.03	0.00	0.15	-0.01	0.40
	17	0.05	0.05	0.02	0.35	0.17	0.43
	19	0.10	0.10	0.00	5.75	4.38	1.81
	21	0.07	0.07	0.00	0.76	0.55	0.53
	23	0.05	0.05	0.01	0.28	0.12	0.37

Table 15: Summary statistics of yield spread and mid yield for European selected bonds.

Country	Maturity(Year) Median	Maturity(Year) Max	Maturity(Year) Min	Issue Date Earliest	Issue Date Latest	Mature Date Earliest	Mature Date Latest	Currency
China								
	5	7	1	14/01/2019	11/08/2021	05/06/2022	02/12/2026	CNY
the US								
	14	60	3	08/02/2019	28/07/2021	15/03/2024	07/10/2079	USD
Europe								
France	6	11	4	29/01/2019	08/10/2020	28/08/2023	26/05/2031	Euro
Germany	6	10	3	04/06/2019	8/12/2020	14/11/2022	07/11/2029	Euro
Spain	6.5	7	5	06/02/2020	18/12/2020	14/11/2022	07/11/2029	Euro
Belgium	6.5	7	6	16/06/2020	10/09/2020	10/09/2026	16/06/2027	Euro

Table 16: Statistics of sample corporate bonds including both green bonds and conventional bonds. The sample number of companies for France, Germany, Spain, Belgium and Italy are 5, 4, 2, and 1 respectively.

Country	Year	Unemployment	GDP in trillion(USD)
the US	2019	3.68	21.00
	2020	8.09	20.00
	2021	5.36	23.00
China	2019	4.52	14.28
	2020	5.00	14.72
	2021	3.96	17.70
Germany	2019	3.14	3.86
	2020	3.81	3.81
	2021	5.80	4.55
Italy	2019	9.95	2.00
	2020	9.16	1.88
	2021	10.20	2.12
Spain	2019	14.10	1.39
	2020	15.53	1.28
	2021	13.33	1.43
Belgium	2019	5.36	0.53
	2020	5.55	0.52
	2021	5.80	0.58
France	2019	8.41	2.73
	2020	8.01	2.62
	2021	8.00	2.94

Table 17: Unemployment and GDP statistics for sample countries from 2019 to 2021.

8.2 Macro statistics

Letter Grade	Grade	Capacity to Repay
AAA	Investment	Extremely
AA+, AA, AA-	Investment	Very strong
A+, A, A-	Investment	Strong
BBB+, BBB, BBB-	Investment	Adequate
BB+, BB	Speculative	Faces major future uncertainties
B	Speculative	Faces major uncertainties
CCC	Speculative	Currently vulnerable
CC	Speculative	Currently highly vulnerable
C	Speculative	Has filed bankruptcy petition
D	Speculative	In default

.

Table 18: S&P rating chart, source: [a]

[a]https://www.spglobal.com/ratings

8.3 Supplementary results: Sensitivity check for endogenous and exogenous variables for the Hausman

8.3.1 Supplementary analysis of Hypothesis 2 treating GDP as endogenous

	China		US		Europe	
VARIABLES	Pooled OLS	HT	Pooled OLS	HT	Pooled OLS	HT
Yld Spread	45.03***	109.7***	10.67***	4.887***	5.926***	0.757***
(TVendogenous)	(0.224)	(3.318)	(0.112)	(0.501)	(0.140)	(0.261)
Bid Yld Change	0.559*	0.326***	0.708***	0.736***	9.14e-05	6.19e-05
(TVexogenous)	(0.299)	(0.0905)	(0.212)	(0.133)	(0.000295)	(0.000143)
Green Bond	-0.141***	-62.83	0.167***	-3.760	-0.0303***	0.530
(TIendogenous)	(0.0184)	(133.8)	(0.00866)	(8.127)	(0.00482)	(0.557)
Maturity	0.288***	-1.096	0.0548***	0.0531*	0.0652***	0.0181
(TIexogenous)	(0.00640)	(4.378)	(0.000271)	(0.0291)	(0.00129)	(0.0462)
GDP	-0.0773***	-0.0824***	0.114***	0.117***	-0.0351***	0.0914***
(TVendogenous)	(0.00681)	(0.00208)	(0.00458)	(0.00288)	(0.00187)	(0.0102)
Constant	0.486***	33.33	-1.232***	1.116	-0.597***	-0.691***
	(0.119)	(78.02)	(0.0804)	(4.180)	(0.0136)	(0.250)
Observations	15,360	15,360	7,380	7,380	7,700	7,700
Number of Bonds	30	30	18	18	20	20
R-squared	0.764		0.866		0.293	

***, ** Statistically significant at the 1% level and 5% level, respectively.

Table 19: Supplementary analysis of HYPOTHESIS 2 with treating GDP as endogenous.

The results show that the change of GDP from exogenous to endogenous does not cause to difference, the results are identical from Table 7.

8.3.2 Supplementary analysis of Hypothesis 3 treating size premium as endogenous

	China			US			Europe		
VARIABLES	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Yld Spread	39.81***	18.35***	39.81***	11.27***	39.94***	11.26***	20.78***	40.83***	20.82***
(TVendogenous)	(0.197)	(0.409)	(0.197)	(0.452)	(1.555)	(0.452)	(0.225)	(0.318)	(0.225)
Bid Yld Change	0.994***	0.952***	0.993***	4.042***	3.670***	4.039***	-0.00126	0.000178	-0.00123
(TVexogenous)	(0.254)	(0.126)	(0.254)	(0.337)	(0.260)	(0.337)	(0.00172)	(0.000707)	(0.00172)
Green Bond	-0.315***	27.31	-0.425*	0.0600***	9.286	0.611	0.233***	38.24	0.00666
(TIendogenous)	(0.0169)	(81.67)	(0.230)	(0.0176)	(16.53)	(0.580)	(0.0223)	(142.0)	(0.0812)
Maturity	0.142***	-0.317	0.142***	0.0594***	0.0651	0.0594***	0.206***	-1.995	0.206***
(TIexogenous)	(0.00619)	(2.678)	(0.00619)	(0.000520)	(0.0618)	(0.000520)	(0.00568)	(8.905)	(0.00568)
Size premium	-1.045***	-0.517***	-1.075***	1.512***	1.362***	1.581***	0.240***	0.206***	0.220***
(TVendogenous)	(0.0637)	(0.0330)	(0.0899)	(0.0724)	(0.0563)	(0.102)	(0.00691)	(0.00289)	(0.00979)
Size premium*GB			0.0606			-0.137			0.0401***
(Interaction term)			(0.127)			(0.144)			(0.0138)
Constant	2.003***	-9.461	2.058***	-5.036***	-10.94	-5.311***	-3.496***	-10.23	-3.387***
	(0.118)	(34.57)	(0.165)	(0.292)	(8.672)	(0.411)	(0.0563)	(21.62)	(0.0677)
Observations		14952			5,340			10,392	
Number of Bonds		24			20			24	

***, ** Statistically significant at the 1% level and 5% level, respectively.

Table 20: Supplementary analysis of HYPOTHESIS 3 with treating size premium as endogenous.