Does Management Matter?

15.034 Econometrics for Managers

Seth Chatterton, Zack Horton, Sasha Lioutikova

Table of Contents:

Introduction

By applying econometric modeling principles to data from the World Management Survey, we can answer the question, "How does the management of a hospital affect patient outcomes?" Through a cross-sectional study of hospitals from 5 countries, surveyed over 3 years, there is a robust and expansive dataset that can be explored. Within this data, the relationship between management quality (`zmanagement`) and acute myocardial infarction (heart attack) mortality rates (`zami_rate`) is of particular interest. Two modeling approaches that are important to consider in this situation are fixed-effects and instrumental variable estimation. Through these approaches, we will be able to uncover the causal relationship between management practices on heart attack mortality rates.

Selection of Controls

Before modeling begins, it is important to understand the underlying relationships in the data to adequately select control variables to reduce any endogeneity concerns and limit overcontrolling. One key aspect of a good control variable is the relationship between the "`D`" and "`Y`" variables; controls should be influential to heart attack mortality rates (`Y`) but unrelated to management practices (`D`) to not over-control. Therefore, variables like `management`, `lmba`, and `mba` were determined to be inappropriate controls, as their inclusion would change the intended question and overcontrol for management. However, variables like `hos_bed`, `hos_numcompetitors`, `hos_fprofit`, and `hos_nfprofit` are influential in patient outcomes but will not overcontrol for management practices. For these variables, the DAG (directed acyclic graph) would be similar to that seen in Appendix A: Figure 1. We additionally control for `grid_temp_new`, the average temperature of the location of the hospital, as research has shown that higher temperatures can increase the rate of heart attacks [1]. Other geographic variables are not related to heart attacks or management and are not included.

One unique aspect of this data is that it was collected from hospitals through a survey. Therefore, it is important to consider response errors and the reliability of the respondent, as that may affect the underlying data. Thankfully, the World Management Survey has quantified this in the variable `survey_reliability`. By controlling for this, the models consider the reliability of responses. The World Management Survey data also includes another variable, `survey_reliability_miss`, which indicates if `survey_reliability` is unknown. It is important to control for instances of missingness as hospitals that have unknown measures of reliability (where `survey_reliability` is missing) may be different from those where reliability is known. Now appropriate controls have been selected, modeling can begin.

Fixed-Effect Modeling

Fixed-Effects (FE) models uniquely include the "`D`" variable, the "`X`" controls, and another set of terms for the "categorical fixed effect". One way to understand this "categorical fixed effect" is to add a dummy variable for each value of the variable. In this particular example, the FE model includes country fixed effects, which is equivalent to adding a dummy variable for each value within the country variable and running a typical OLS regression with the same "`D`" variable and "`X`" controls.

Country fixed effects were chosen due its increased precision (improved standard error, statistical significance for variables, and increased significance for the global F-test) compared to when the model was run with `region_survey` as the fixed effect. These two variables are natural choices for an FE model, as it is easy to believe that the quality of management and health may be different between countries (see Appendix B: Figure 1 and Appendix C: Table 1). One variable that was omitted from the control was `yy` (and its other forms, `yy06` and `yy09`).

This decision was made because each country was exclusively surveyed within a single year, except for the United Kingdom (see Appendix B: Figures 3, 4, and 5). To add confidence in this decision, FE models with and without the `yy09` variable were created and had negligible differences between the two. When year indicator variables were included, models saw increased standard errors and less significance at both the variable level and for the global F-test. We did not include `hospital_id`, `com_id`, or `analyst` because it would be difficult to find within-variation for any of these groups.

Instrumental Variables Estimation

We also created an instrumental variable model, which we can use if we believe that `zmanagement` is endogenous. This might be the case since we are using survey data which may have a selected sample bias, or have measurement error of the management scores. We do not know the methodology used to pick these particular hospitals and survey respondents. We also do not know how well `zmanagement` represents the actual management abilities of a hospital, since there are so many complex factors that go into this one number.

In our analysis, we use variables related to nearby combined Medical and Business schools. We use `logcom_ttime`, `com_lage`, and `com_lage_miss` as our instruments. The other Medical-Business school variables are related to these and will not add much more analytical power. The first variable, `logcom_ttime`, is the natural log of the commuting time to the nearest combined Medical and Business (M-B) school. We use this variable with the hypothesis that hospitals closer to an M-B school will have more graduates from that school go there. We think that having more M-B graduates will increase a hospital's `zmanagement` score (see Appendix B: Figure 6). The next instrument is `com_lage` and `com_lage_miss`, which represent the natural log of the age of the joint M-B and whether that data is missing. The hypothesis behind including these variables as instruments is that more established schools will have had more time to bring their expertise to their community (see Appendix B: Figure 7). See Appendix C: Figure 3 for the coefficients of stage 1 of the instrumental variable model.

Findings & Concerns

To quantify management quality, we z-score normalized the `management` variable (creating `zmanagement`). This was done to increase the interpretability of causal estimates, as now, results can be interpreted as, “How does increasing management quality by 1 standard deviation of ‘average quality’ affect patient outcomes?” The same was done for heart attack mortality rate (creating `zami_rate`), so coefficients can be interpreted as, “Increasing variable `X` by 1 changes heart attack mortality rate by `X` coefficient standard deviations for `zami_rate`.”

Through a robust modeling process, our fixed effect model provided us with interesting results. In Appendix C: Table 2, we can see that it is estimated that by increasing `zmanagment` by 1 standard deviation (relative to country) we can decrease the `zami_rate` by 0.1824 standard deviations (relative to country). This follows our logic that improving management practices has a positive influence on patient outcomes. Additionally, we find that the coefficients related to the control variables also follow our initial hypotheses. Mainly, we see that `grid_temp_new` has a slight increase in the mortality rate (though not significant) and that when survey reliability is missing (`survey_reliability_miss`), it is estimated to increase heart attack mortality by 0.2 standard deviations. Overall, we are happy with the outcome of this approach and feel that we have been able to appropriately model the causal relationship between quality of management and heart attack mortality rates.

We find a similar conclusion from our instrumental variable model. Appendix C: Table 4, shows that increasing `zmanagment` by 1 decreases `zami_rate` by 1.0456 standard deviations, all else being equal. This is a larger coefficient than the fixed effect model. However, the model has an f-statistic p-value of 0.4013, meaning we can't draw any firm conclusions. This can be explained by the fact that there is only a small amount of variation of `zmanagement` explained by the instrumental variables, so the model is working with less information.

Overall, we took a robust and methodological approach to our causal modeling. However, omitted variables and general endogeneity are always of top concern in situations like this. Luckily, we do not believe any inclusion variable bias is present here. For OVB and endogeneity, one issue may be related to data collection and some "unique" aspects of the hospitals that were surveyed that are unobserved within our data. Additionally, we do not know the capabilities or specialties of each hospital, and this could introduce omitted variables not properly capture the ways hospitals are different outside of the variables we did include. Finally, there is the slight potential for overcontrolling by including something like `hos_num_competitors` because of its relationship with patients coming to a hospital when they have fewer alternatives available. However, we included it because the superior management of a hospital may make it hard for other hospitals to survive, or the presence of high competition may drive management to be better. Therefore, we would rather run the risk of overcontrolling and avoid the endogeneity.

Conclusion

In conclusion, we think that management does matter in hospital settings. Controlling for many relevant factors including country fixed effects, hospital type, size of the hospital, survey quality, and temperature, both of our models showed a negative correlation between the quality of hospital management and the rate of heart attack mortality. One potential causal mechanism is that better management leads to better organization and knowledge sharing between hospital staff, leading to more prompt responses in life-saving scenarios. Future work could go into further depth to explore these mechanisms.

Works Cited

[1] “As temperatures rise, so does the risk of heart problems”.

https://about.kaiserpermanente.org/health-and-wellness/health-tips/as-temperatures-rise-so-does-the-risk-of-heart-problems

Appendix A: Causal DAG

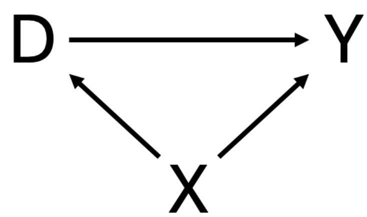

***Figure 1:** A causal relationship where the control, X, influences the variable of interest, D, and the outcome of interest, Y. Here, it **is** appropriate to control for X*

Appendix B: EDA

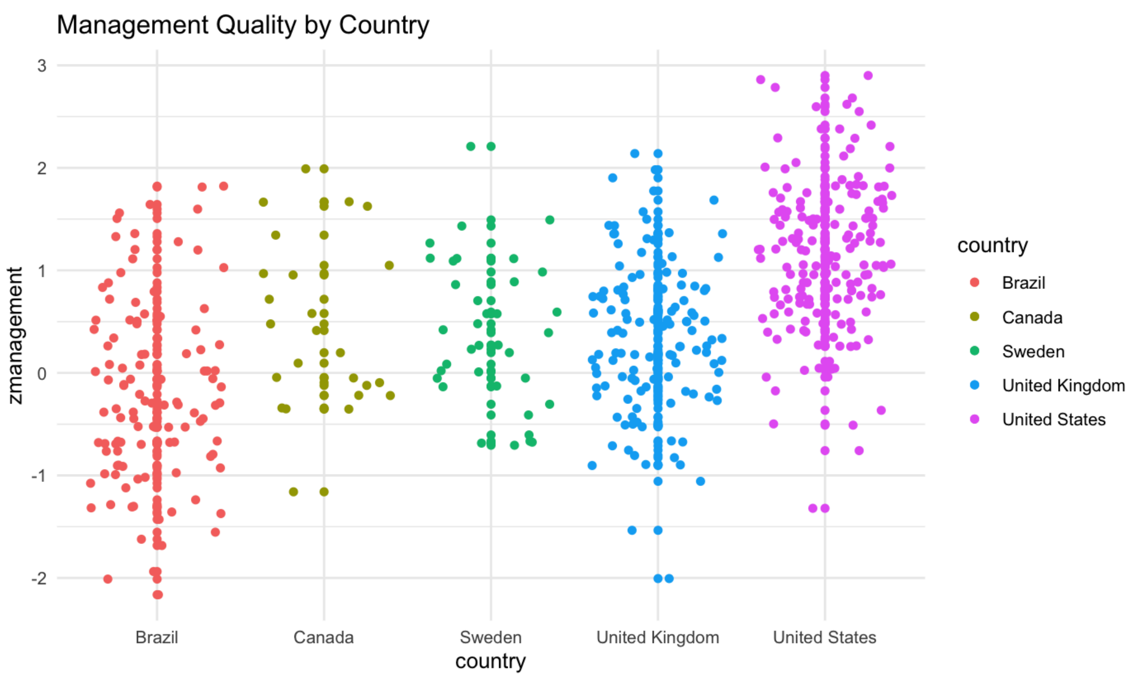

***Figure 1:** Showing that management quality differs across countries, see Appendix C: Table X for more verbose results with a simple linear regression*

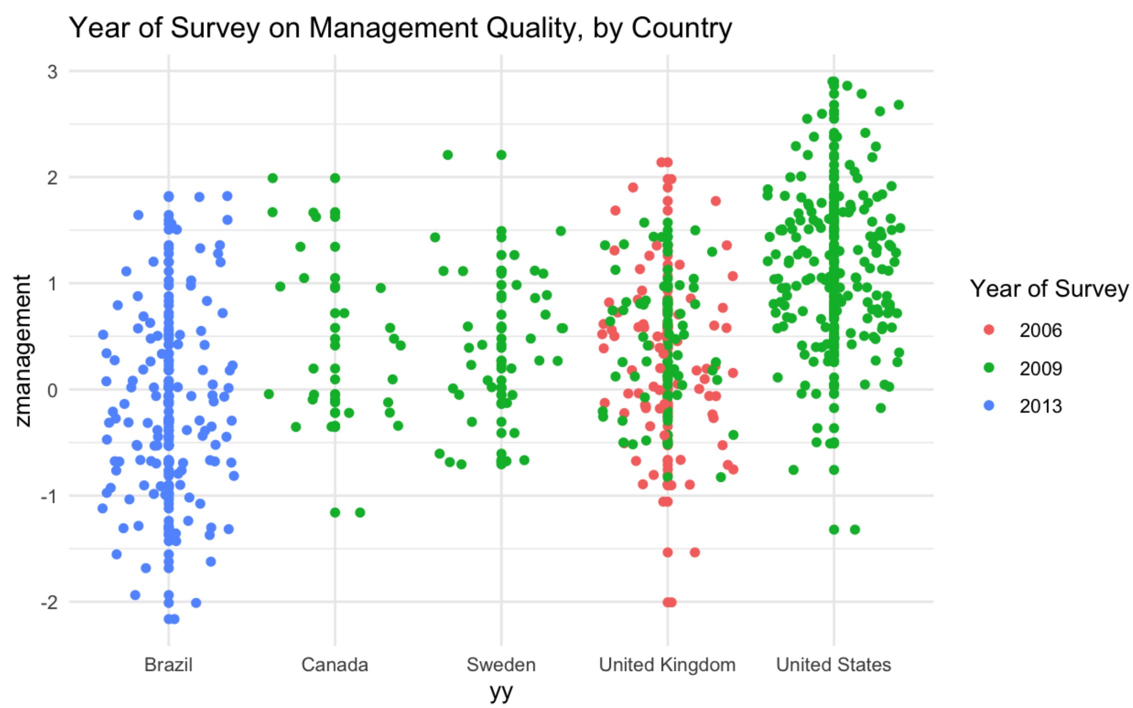

Figure 2: *Showing that management differs across countries and has limited change overtime*

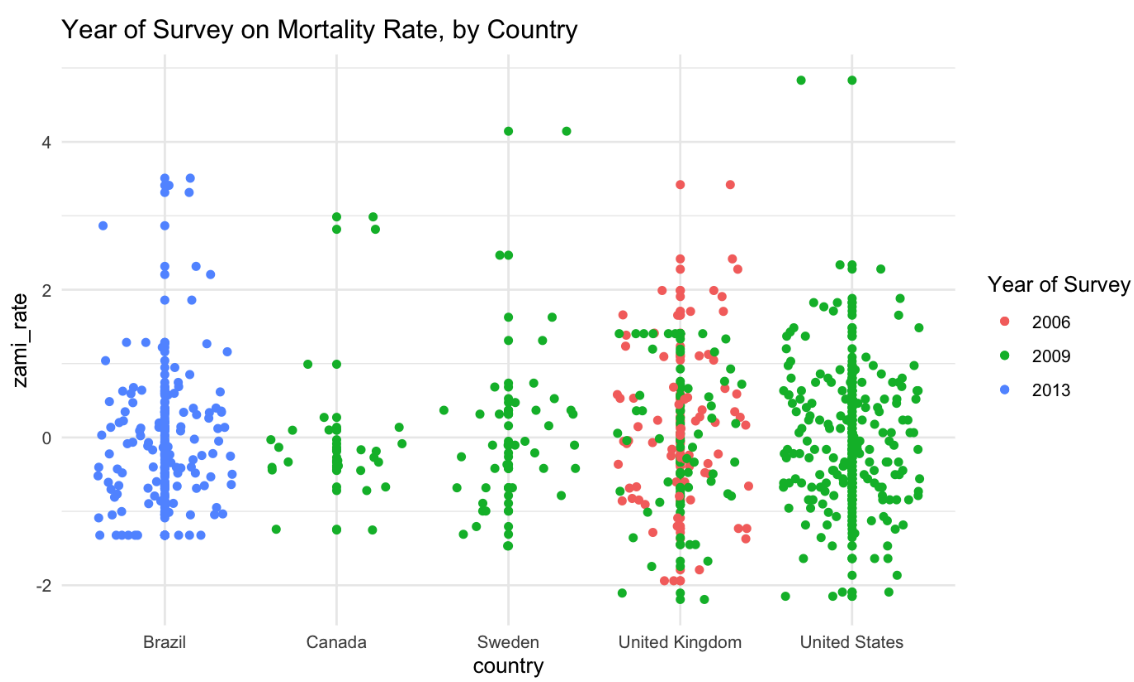

Figure 3: *Showing the year that each country was surveyed*

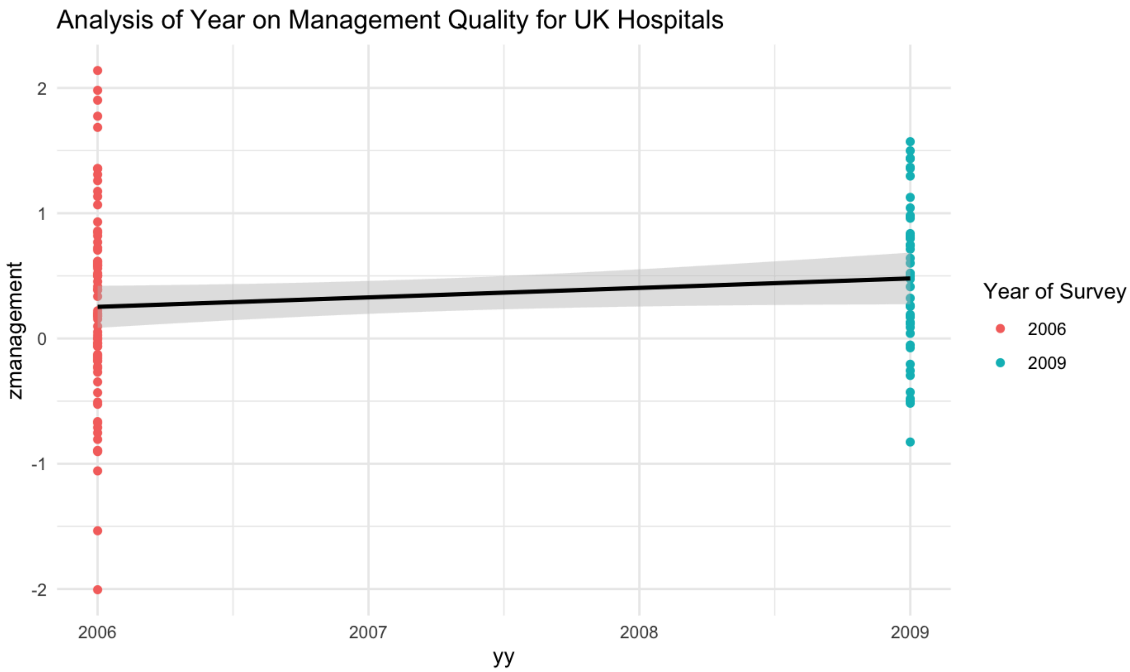

Figure 4: *Illustrating that no significant relationship exists between "yy" and "zmanagement" for UK Hospitals when comparing survey results from 2006 to 2009*

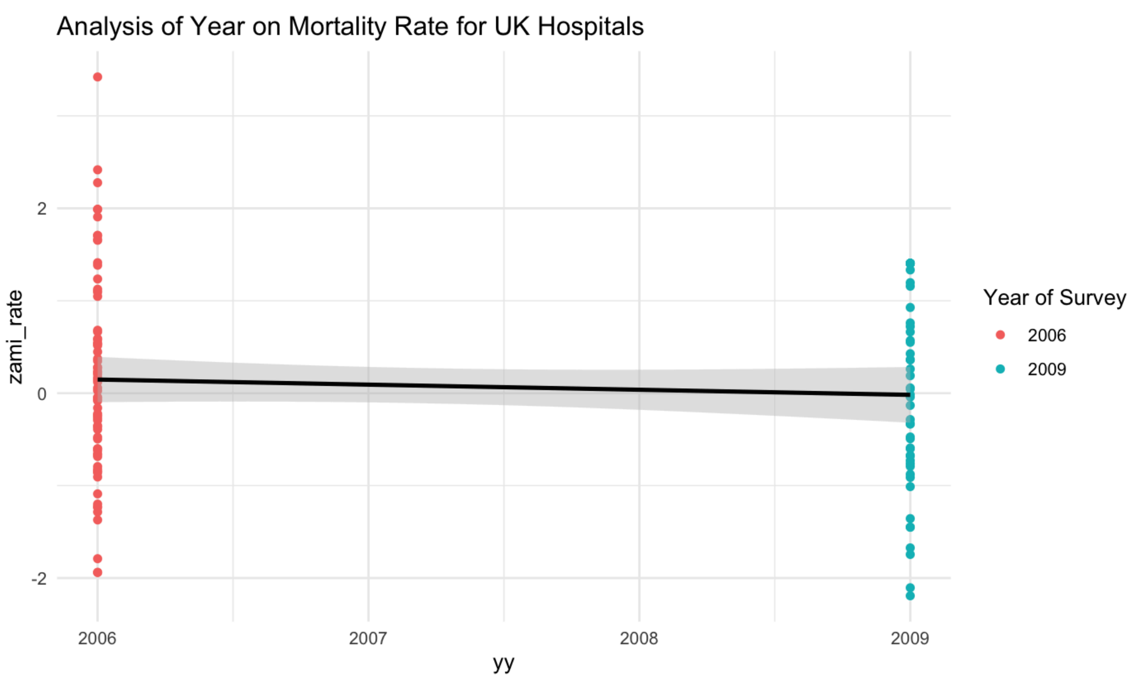

Figure 5: *Illustrating that no significant relationship exists between "yy" and "zami_rate" for UK Hospitals when comparing survey results from 2006 to 2009*

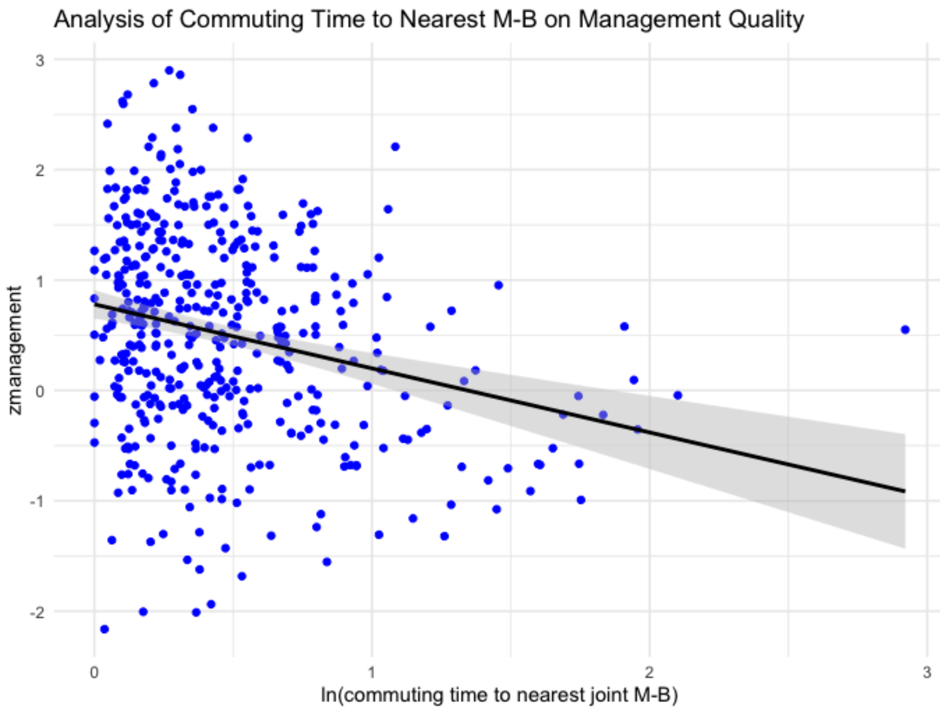

Figure 6: *Showing that the log of the commuting time to the nearest joint Medical and Business school is potentially related to management scores.*

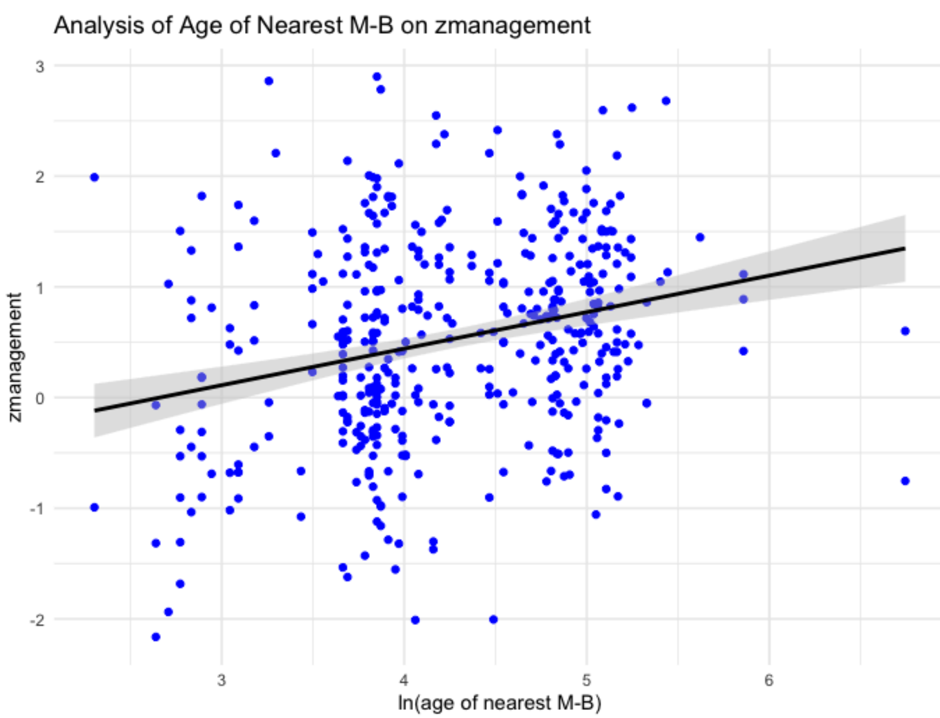

Figure 7: *Showing that the log of the age of the nearest joint Medical and Business school is potentially related to management scores.*

Appendix C: Results & Estimates

Country	Coefficient	Standard Error	P-Value
Intercept [Brazil]	-0.186	0.0743	0.0125
Canada	0.636	0.1750	0.0003
Sweden	0.575	0.1492	0.0001
United Kingdom	0.529	0.1021	3.3e-7
United States	1.33	0.0963	<2e-16

Table 1: *Coefficients when predicting management quality as a function of country lm(zmanagement ~ country), see Appendix B: Figure 1 for related visualization*

Variable	Coefficient	Standard Error	P-Value
zmanagement	-0.1824	0.06488	0.00516
hos_lbed	0.0087	0.00978	0.37461
hos_fprofit	-0.0351	0.21729	0.87201
hos_nfprofit	-0.2665	0.13385	0.04711
hos_numcompetitors	-0.1994	0.08638	0.16756
survey_reliability	0.0039	0.02651	0.88452
survey_reliability_miss	0.1929	2.85211	0.94610
grid_temp_new	0.0095	0.01496	0.52632
grid_temp_new_miss	0.6965	1.62950	0.66927

Table 2: *Summary statistics of final fixed-effects model in R plm(zami_rate ~ zmanagement + hos_lbed + hos_fprofit + hos_nfprofit + hos_numcompetitors + survey_reliability + survey_reliability_miss + grid_temp_new + grid_temp_new_miss, model="within", index="country")*

Variable	Coefficient	Standard Error	P-Value
logcom_ttime	-0.313476	0.099266	0.00170
com_lage	-0.007723	0.059524	0.89683
com_lage_miss	-0.825983	6.122994	0.89275

Table 3: *Coefficients for stage 1 of the instrumental variables model. Controls are included in the model, but not shown in the table for brevity.*
stage_1_lm <- lm(zmanagement ~ logcom_ttime + com_lage + com_lage_miss + as.factor(country) + hos_lbed + hos_fprofit + hos_nfprofit + hos_numcompetitors + survey_reliability + survey_reliability_miss + grid_temp_new + grid_temp_new_miss)

Variable	Coefficient	Standard Error	P-Value
zmanagement	-1.043021	0.513398	0.0428
Canada	0.885774	0.635189	0.1639
Sweden	0.873678	0.609245	0.1523
United Kingdom	0.946296	0.561131	0.0924
United States	1.524804	0.783257	0.0522
hos_lbed	0.005460	0.011728	0.6418
hos_fprofit	0.265066	0.312200	0.3963
hos_nfprofit	-0.075137	0.194510	0.6995
hos_numcompetitors	-0.017080	0.118724	0.8857
survey_reliability	0.104658	0.067228	0.1202
survey_reliability_miss	10.998751	7.212318	0.1280
grid_temp_new	0.007524	0.017750	0.6719
grid_temp_new_miss	0.330931	1.940643	0.8647

Table 4: *Summary statistics of the final instrumental variables model in R*
reg_iv <- ivreg(zami_rate ~ zmanagement + as.factor(country) + hos_lbed + hos_fprofit + hos_nfprofit + hos_numcompetitors + survey_reliability + survey_reliability_miss + grid_temp_new + grid_temp_new_miss
| logcom_ttime + com_lage + com_lage_miss + as.factor(country) + hos_lbed + hos_fprofit + hos_nfprofit + hos_numcompetitors + survey_reliability + survey_reliability_miss + grid_temp_new + grid_temp_new_miss,
data=data)